Thèse de Doctorat

Mention Mathématiques

présentée à *l'École Doctorale en Sciences Technologie et Santé (ED 585)*

de l'Université de Picardie Jules Verne

par

Yohan Hosten

pour obtenir le grade de Docteur de l'Université de Picardie Jules Verne

Sur des propriétés stochastiques de la variation de la somme des chiffres en additionnant un entier fixé

Soutenue le 9 Décembre 2022, après avis des rapporteurs, devant le jury d'examen

M^me^ DAJANI Karma	**PR, Utrecht University**	**Examinatrice**
M. DRMOTA Michael	**PR, Technische Universität Wien**	**Rapporteur**
M. DURAND Fabien	**PR, Université de Picardie Jules Verne**	**Président**
M. DURIEU Olivier	**MCF, Université de Tours**	**Examinateur**
M. HUBERT Pascal	**PR, Université d'Aix Marseille**	**Rapporteur**
M^me^ JANVRESSE Élise	**PR, Université de Picardie Jules Verne**	**Directrice**
M. de la RUE Thierry	**CR CNRS, Université de Rouen Normandie**	**Directeur**

Pour Julia, Margaux, Aubin et la tribu complétée.

Remerciements

Ça y est, c'est le moment. Après plusieurs mois le pied au plancher sur l'autoroute de la rédaction, dans les courbes sinueuses des voies administratives, il est temps de lever le pied, de regarder dans son rétroviseur. Durant ces dernières années, j'ai eu la chance de croiser la route de certains autres voyageurs remarquables et je vais tâcher, dans les lignes qui suivent, d'adresser mes sincères remerciements à chacun d'entre eux.

Je vais commencer par mes membres de jury : je remercie chaleureusement Karma Dajani, Michael Drmota, Fabien Durand, Olivier Durieu et Pascal Hubert d'avoir bien voulu de m'accompagner dans cette dernière étape de la thèse qu'est la soutenance. En particulier, je veux remercier Pascal Hubert que j'ai eu la chance de rencontrer il y a quelques mois et qui m'a accordé un accueil au CIRM qui restera longtemps gravé dans ma mémoire. Je veux remercier également Michael Drmota d'avoir accepté, avec Pascal Hubert, la responsabilité de rapporter cette thèse. Je vous avouerai avoir ressenti beaucoup de fierté en lisant vos rapports.

J'ai toujours dit et je dirai toujours que la qualité de mon travail n'est que le simple reflet de la qualité de mon encadrement. Je remercie donc, avec beaucoup d'émotions, Thierry et Élise, mes deux directeurs. Il n'y a pas de mot pour décrire l'admiration que j'ai pour chacun de vous. Quand tu es venue me voir, Élise (ou Mme Janvresse à l'époque), pour me proposer un sujet il y a quelques années, je n'ai absolument pas hésité. Ayant déjà travaillé avec toi, je savais où j'allais. Et, quand je t'ai rencontré pour la première fois, Thierry, ta bienveillance a annihilé tout début d'appréhension que j'aurais pu avoir. Thierry, j'imagine que tu as souvent entendu/lu que tu étais un très bon directeur de thèse. Aujourd'hui, à l'heure du bilan, je ne peux que renouveler tous les compliments qu'on a pu te faire. Pour Élise, j'ai cette opportunité d'être le premier à pouvoir te le dire : tu es une excellente directrice de thèse et je te souhaite, si c'est ce que tu souhaites également bien sûr, d'avoir beaucoup d'autres doctorants ou doctorantes à l'avenir. Vous m'avez appris cette recherche ardue de l'équilibre entre la justesse d'écriture, la simplicité de l'argumentation et la rigueur mathématique. Il me faudrait bien plus de 180 secondes pour énoncer vos qualités. Mais, en particulier, j'ai pu profiter de votre calme, de votre sérénité face à un doctorant qui pouvait être en perte de vitesse... je n'ose pas écrire "qui pouvait être entêté parfois aussi" !

Je vais maintenant passer à la famille des doctorants/récents docteurs d'Amiens. Comment ne pas commencer par la "Team Thiérache" : Alice et Clément ?! Outre vos aptitudes évidentes en maths, j'admire la patience que vous avez eu envers moi. 8 ans à me supporter (les blagues nulles, le caractère, ...), c'est un vrai challenge : demander à Ludivine. Depuis la L3, on en a fait de la route ensemble et il y a une partie triste en moi à l'idée que nos chemins s'écarteront bientôt. À la même génération doctorale, je remercie Jihade, ma cobureau avec Alice, pour le sourire que tu arrives à transmettre chaque matin. Je remercie ensuite Ismaïl pour les diverses discussions mathématiques (ou pas) dans mon bureau (ou ailleurs) et les moments de détente. Je

remercie les doctorantes de la “Team Philippines”, à savoir Arrrrrianne, Sheila, Cheryl et Karen. En particulier, je me permets de saluer le grand courage dont Cheryl a dû faire preuve pour venir en France. De l'autre côté de la planète, je remercie nos doctorants chiliens Sebastian, Christopher, Felipe et Bastian. Sans ce passage en doctorat, je pense que je n'aurai jamais rencontré autant de personnes d'horizons géographiques différents. Merci à Marouan qui m'a accueilli, à bras ouverts, dans mon bureau le premier jour et avec qui j'ai eu de sacrés discussions sur tout et, plus souvent, sur n'importe quoi. Merci à mon (faussement) grincheux préféré, Henry. Tes analyses footballistiques les lendemains de matchs de Ligue des Champions ou de matchs de sélections me manqueront. J'adresse mes salutations les plus respectueuses aux frères Garnier : Owen et Arthur, nos Bernoulli à la française, dont les vitesses d'analyse (un comble pour des algébristes) et de réflexions me surprendront toujours. Merci également à Gauthier pour les blagounettes et le côté relax dans notre monde de fous, merci à Valérie pour les discussions pédagogiques, Guillaume pour son (ses ?) incroyables histoires de vie et Ahmad pour sa gentillesse. Enfin, je souhaite à la Relève : Nicolas, Maxime et Michael, de prendre bien soin de cette belle maison qu'est le Bureau des Doctorants et de vivre, pendant les prochaines années, de folles aventures et d'aussi belles rencontres que celles dont j'ai pu profiter. À chacun et chacune, je souhaite sincèrement le meilleur pour vos vies aussi bien professionnellement que personnellement. (Mon seul regret de doctorant est de ne pas avoir réussi à organiser un match de foot contre les permanents. Quelqu'un se dévoue ?)

La famille des doctorants fait partie de la grande famille du LAMFA, je me servirai de ce paragraphe pour adresser un mot à ses membres. Tout d'abord, je commencerai par ceux qui font partis (et j'aimerai assez que cela change) des “invisibles de la recherche”. Ceux sans qui nous, chercheurs, ne serions rien. Ils représentent notre lien avec les institutions et nous aident (voire plus) quand nous devons partir en déplacement, imprimer nos thèses, réparer nos ordinateurs, installer nos bureaux, $\cdots$ J'adresse donc mes remerciements à Isabelle, Mylène, Christelle, Laurent et Étienne pour toute l'aide apportée (avec, toujours, le sourire) au cours des dernières années. Merci ensuite à notre directeur de laboratoire, David, pour toute sa gentillesse et son efficacité. Merci aussi à Ivan, notre ancien directeur, pour l'écoute quand j'en ai eu besoin. Au sein de mon équipe de recherche, j'adresse mes remerciements à Martin, Samuel, Benoît, Fred, Clémence, Gabriel et Aihua. De façon plus générale, merci à Vivien, Alberto, Hervé, Ramla, Karine, Daniel, Jean-Paul, Yann, Sungsoon, Radu, Moussa, Stéphane et Louis. J'ai une pensée pour trois anciens du LAMFA également : Marion, Olivier et Youcef. Tous et toutes, que ce fût quand j'étais votre étudiant ou dans la salle café, je me souviendrai de vous avec bonheur.

Sur un volet plus personnel, je remercie mes parents pour leur éducation. On dit qu'il n'y a pas de bons parents, que des gens qui font de leur mieux. Aujourd'hui, depuis la naissance de Bébé Julia, j'en suis pleinement conscient. J'imagine les sacrifices que vous avez dû faire pour que vos enfants puissent avoir ce qu'ils ont aujourd'hui. J'espère avoir trouvé, par le travail, une forme de remerciements adaptée. Merci à ma soeur Gaëlle (ou “la fille de ma mère et de mon père” ?) pour les différentes formes de soutien. Tu as traversé vaillamment des épreuves que l'on ne souhaiterait pas à son pire ennemi, je ne suis pas sûr d'avoir cette force de mental ou de caractère. Les taquineries depuis tout petit n'étaient peut-être que la traduction d'une forme de (bonne) jalousie après tout ? Merci à Jérôme et Margaux ainsi qu'à ma belle-famille, Bruno, Yaël et Maxime pour les parenthèses heureuses. J'ai une pensée pour mon parrain et mon grand-père maternel que j'aurai aimé rendre fier un peu plus longtemps. Merci à ma conjointe Ludivine : que de changements nous avons vécu tous les deux pendant ces 3 ans. Je ne te promets pas d'en connaître autant à l'avenir mais je resterai le plus longtemps possible à tes côtés pour danser sous la pluie (et rire sous le soleil évidemment).

J'adresse un petit mot à mes compagnons de route en commençant par mon plus vieil ami Hugo. De nos 3 ans jusqu'à aujourd'hui et demain encore, nos blagues évolueront (pour le meilleur et pour le pire !) mais on en rigolera toujours autant, merci pour cela. Je te souhaite beaucoup de bonheur avec Blanche (que je remercie ici aussi). Je remercie également mon second plus vieil ami, Samuel et m'excuse pour toutes les fois où j'ai pu l'embêter étant petit (même si c'était, bien évidemment, Hugo le cerveau de ces machinations !). Merci ensuite à Manon, Florian, Valentin et Marjojo, mes fidèles camarades de lycée, pour tous ses moments vécus et à vivre encore longtemps je l'espère (je ne suis, par contre, pas certain d'être encore assez souple pour des Twister's après 22h). Merci à Julian, mon armoire à glace préféré. J'espère un jour avoir tes talents de narration (et pouvoir te mettre une raclée au béhourd mais, là, je rêve beaucoup !). Merci à Justine également pour nos discussions, entre autre, "du terrain" et me remettre un peu les pieds sur terre. Le temps passe et passe et passe et beaucoup de choses ont changé et, pourtant vous faites partis de mes constantes de vie. Beau compliment de la part d'un matheux, non ?

Pêle-mêle, je remercie M. Cherkaoui, mon premier capitaine mathématique et Mme. Spriet pour être, tous les deux, une source d'inspiration continuelle. Merci à M. Charifou et Mme Daniels pour avoir vu en moi ce que je ne voyais pas. Merci à mes Youtubers préférés (BBoy et Bob Lennon) pour être une source de divertissements (et de blagues) inépuisables. Merci aux films, aux livres et à la musique d'exister et de me permettre d'avoir un échappatoire au maelström qui peut se dérouler dans la tête d'un thésard.

Mes derniers mots ici seront pour ma fille, la prunelle de mes yeux, mon petit soleil. Tu ne sais pas encore parler, encore moins lire, mais le temps me donnera vite tort. Peut-être un jour liras-tu ces mots (et, j'en serai très honoré, ce travail). Alors tu sauras que tu as changé, plus que quiconque, mon monde et que pas un jour ne passe sans que je sois fier de toi et de ton évolution. Ces mots ne sont que ceux d'un être humain sur plusieurs millards mais ce sont les mots de la personne qui te chérit le plus au monde (avec ta mère évidemment).

Votre dévoué ami/collègue/(beau-)frère/(beau-)fils/papa/conjoint,

Yohan.

Table des matières

Table des figures

Introduction générale

Ce travail de thèse porte sur un problème de numération dépendant, par nature, de la manière d'écrire les nombres. Cela amène à une première question mathématique.

Comment écrire les nombres ?

La réponse n'est pas si simple. Historiquement déjà, c'est très compliqué. La réponse ferait intervenir beaucoup de civilisations différentes : égyptienne, grecque, chinoise,$\cdots$ Et, bien qu'étant passionnante, cette histoire sera probablement mal rapportée par un étudiant en mathématiques qui écrirait son introduction de thèse. Mathématiquement par contre, l'exercice sera plus simple. En fait, il existe plusieurs façons d'écrire les nombres. La plus usuelle est, du fait de nos dix doigts, le *système de numération positionnel décimal*. Son principe provient de la propriété que, pour tout entier naturel non nul n, il existe une unique suite de poids $(n_k)_{k\geq 0}$ dans l'ensemble $\{0, 1, \cdots, 9\}$, appelés *chiffres*, telle que

$$n = \sum_{k\geq 0} n_k 10^k.$$

Bien entendu, il n'y a qu'un nombre fini de poids non-nuls. On note donc $\ell := \max\{i \ : \ n_i \neq 0\}$ et ainsi réduire la formule plus haut en

$$n = n_\ell \times 10^\ell + \cdots + n_1 \times 10^1 + n_0.$$

Il en découle l'écriture dite en *base* 10 (ou *décimale*) de n

$$\overline{n_\ell \cdots n_0}^{10} := n = \sum_{k=0}^{\ell} n_k 10^k.$$

Dans l'expression *système de numération positionnel décimal*, le mot *décimal* provient du fait qu'il y a 10 chiffres possibles et le mot *positionnel* du fait que la valeur d'un chiffre dépend de la position de celui-ci dans le nombre (il y a le chiffre des *unités*, celui des *dizaines*, celui des *centaines*, $\cdots$).

Le mathématicien aime bien effacer les particularités des exemples qu'il choisit pour en tirer l'essence d'un problème plus général de sorte à y révéler des structures dissimulées sous leur surface. Après tout, pourquoi donner autant d'importance au nombre 10 ? En effet, si l'être humain avait eu 12 doigts, il aurait alors imaginé un système d'écriture adapté à ses 12 doigts et, cela, aussi naturellement que, nous, nous avons imaginé notre système décimal. Le nombre 10 n'a pas plus d'importance que le nombre 12. Ainsi, on peut généraliser ce système d'écriture décimal en un *système de numération positionnel en base entière*. Fixons un entier $b \geq 2$, ce sera notre base de

numération. Le principe est le même qu'en base 10 : tout entier naturel non nul n peut s'écrire à l'aide d'une unique suite de poids $(n_k)_{k\geq 0}$ dans l'ensemble $\{0, 1, \cdots, b-1\}$, que nous appellerons encore *chiffres*, telle que

$$n = \sum_{k\geq 0} n_k b^k.$$

Comme tout à l'heure, il n'existe qu'un nombre fini de poids strictement positifs, on note ℓ l'indice maximal du dernier chiffre non nul et on écrit n en *base entière* comme

$$\overline{n_\ell \cdots n_0}^b := n.$$

On a vu que le cas $b = 10$ est le plus usité dans le monde. Le cas $b = 2$ est son concurrent direct : c'est le *système de numération positionnel binaire* (ou *binaire* simplement). En effet, c'est le langage des ordinateurs où, par "ordinateurs", on entend "tout appareil qui contient une unité centrale de traitement".

On peut aller encore plus loin dans la généralisation ! En effet, jusqu'à présent, on a décomposé les entiers selon une suite de puissances en pondérant chaque terme par un chiffre de la base (entière) b. Mais, après tout, pourquoi une suite de puissances ? Pourquoi ne pas décomposer les entiers selon une simple suite de nombres entiers ? Et bien, parce que ça ne marche pas toujours. Une suite permettant de décomposer les nombres entiers en une somme de ces termes est appelé une suite *complète*. Mais être complète n'est pas suffisant si l'on souhaite (et c'est plutôt pratique) l'unicité de l'écriture. On peut l'obtenir en ajoutant des contraintes à l'éventuelle décomposition. Voici un exemple d'une telle généralisation. Elle fait entrer en scène une suite qui est très probablement la plus connue du monde mathématique : la *suite de Fibonacci*. Elle est définie comme suit

$$\begin{cases} F_1 &= 1, \\ F_2 &= 1, \\ F_k &= F_{k-1} + F_{k-2} \quad \text{si } k \geq 3. \end{cases}$$

À partir de cette suite, on peut effectivement définir un système de numération : tout entier naturel non-nul peut se décomposer comme une somme de termes de la suite de Fibonacci. Cependant, il n'y a pas unicité de la décomposition. Pour l'avoir, il faut, comme on l'a suggéré plus tôt, ajouter une contrainte d'écriture. Ici, on interdira l'utilisation de deux termes consécutifs. Le *théorème de Zeckendorf* [30] dit que pour tout entier naturel non nul n, il existe une unique suite de chiffres $(n_k)_{k\geq 2}$ dans $\{0, 1\}$ ne contenant jamais deux 1 consécutifs et telle que

$$n = \sum_{k\geq 2} n_k F_k.$$

Comme d'habitude, il n'y a en réalité qu'un nombre fini de termes dans cette sommation. Notons encore ℓ l'indice du dernier terme de la suite utilisé. Cela conduit à l'écriture

$$\overline{n_\ell \cdots n_2}^Z := n.$$

On peut penser à d'autres pistes de généralisations comme des systèmes de numérations où b n'est pas un entier mais un nombre réel. Et bien, sous conditions, il est possible d'imaginer de tels systèmes d'écriture mais ils rendront notre problème très différent. Mais d'ailleurs...

Et notre problème dans tout ça ?

Étant donné un entier $r \in \mathbb{N}$, notre problème consiste à étudier la variation de la somme des chiffres lorsqu'on ajoute r à un entier. Plus précisément, en notant $s(n)$ la somme des chiffres de n dans un système d'écriture fixé, on s'intéresse à la suite

$$\Delta^{(r)}(n) := s(n+r) - s(n).$$

Comme annoncé dans les premières lignes de l'introduction, notre problème dépend effectivement de la manière d'écrire les nombres. Dans cette thèse, le premier chapitre porte sur l'étude de cette suite dans le cas des systèmes en base entière et le second dans le cas de la représentation de Zeckendorf. La dépendance des notations au système d'écriture est consciemment omise : on fixera une base entière b dans le premier chapitre et le second chapitre sera traité indépendamment du premier.

L'intérêt porté à cette suite vient du fait que cette variation est intrinsèquement liée au comportement de l'addition dans le système de numération considéré. En effet, en base $b \geq 2$, si on note c le nombre de retenues créées pendant l'addition $n + r$ (en base b) alors

$$\Delta^{(r)}(n) = s(r) - c(b-1).$$

En binaire, l'étude de cette suite a beaucoup d'intérêt dans la compréhension des erreurs qu'une unité centrale de traitement pourrait faire.

Toute étude de suites devrait commencer par le calcul des premiers termes. Sur la Figure 1, on a tracé quelques exemples pour divers systèmes de numérations.

Visuellement, on peut constater que les valeurs prises par $\Delta^{(r)}$ sont réparties d'une manière plus ou moins régulière. On a envie de dire que $\Delta^{(r)}$ prend une valeur donnée pour une certaine proportion d'entiers. Une question naturelle (a priori mal posée) est donc

à quel point $\Delta^{(r)}$ peut prendre une valeur donnée ?

Dans le cas où la base est un nombre entier, on remarque que chaque ligne de niveau présente une périodicité : en fait, Bésineau a montré dans [3] que chaque ligne contient un nombre fini de suites arithmétiques. Il démontre que, dans le cas de la base entière, pour tout entier relatif $d \in \mathbb{Z}$, la limite

$$\lim_{N \to +\infty} \frac{1}{N} |\{n < N \ : \ \Delta^{(r)}(n) = d\}|$$

existe. Cette limite est la *densité asymptotique* de l'ensemble $\{n \in \mathbb{N} \ : \ \Delta^{(r)}(n) = d\}$. On la note $\mu^{(r)}(d)$. Ainsi, Bésineau a montré qu'en base entière, la question plus haut est bien posée. Comme souvent en recherche, répondre à une question pousse à s'en poser (au moins) une nouvelle... ici on se demande assez naturellement combien peut valoir $\mu^{(r)}(d)$? Bésineau a alors poussé son étude et est parvenu à la formule suivante : si k est un chiffre en base $b \geq 2$ alors, pour tout $d \in \mathbb{Z}$

$$\mu^{(br+k)}(d) = \frac{b-k}{b}\mu^{(r)}(d-k) + \frac{k}{b}\mu^{(r+1)}(d+b-k). \tag{1}$$

Il s'agit d'une relation de récurrence sur la famille de mesures $(\mu^{(r)})_{r \in \mathbb{N}}$ où, à chaque application de celle-ci, on fait baisser le nombre de chiffres de r. Il suffit donc de connaître les mesures $\mu^{(r)}$ pour les entiers r dont l'écriture en base b ne contient qu'un chiffre pour en déduire toutes les

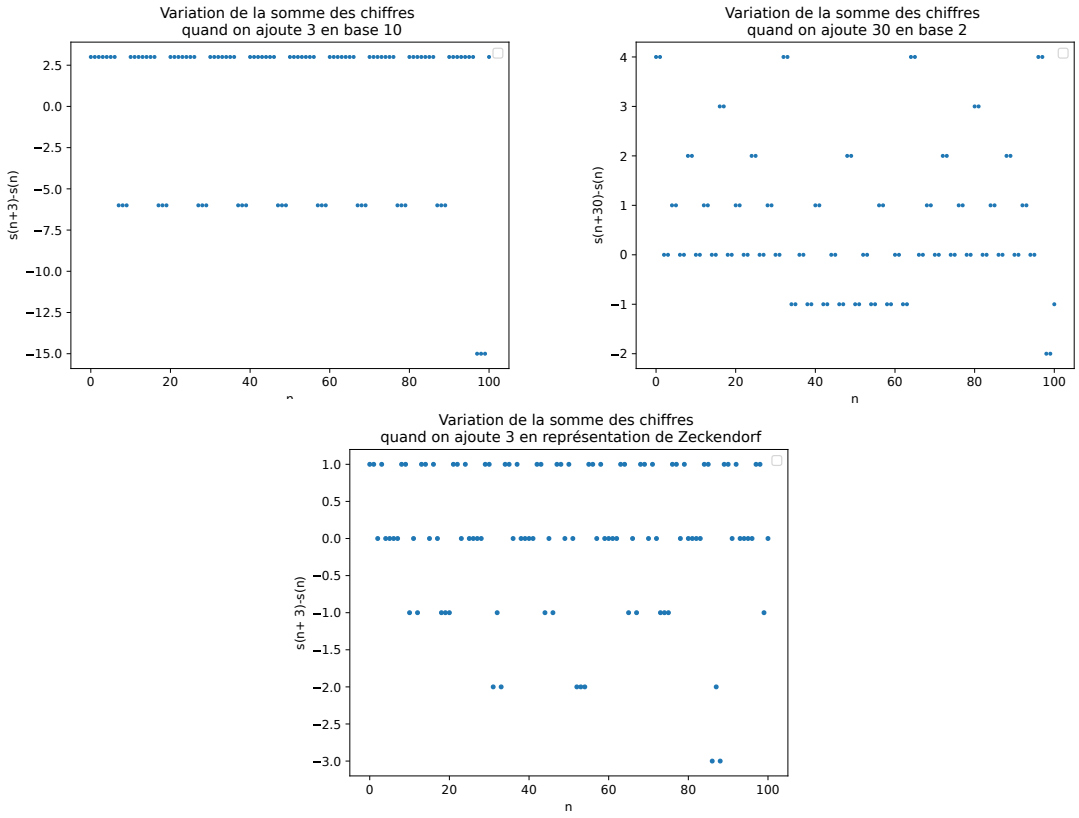


Figure 1: Quelques exemples de tracé de $\Delta^{(r)}$ avec, de gauche à droite $r = 3$ en base 10, $r = 30$ en base 2 et $r = 3$ en représentation de Zeckendorf

autres mesures. En réalité, il suffit de ne connaître que $\mu^{(0)}$ et $\mu^{(1)}$. En calculant ces deux mesures, Bésineau s'est rendu compte que celles-ci étaient des mesures de probabilité et donc, par (1.1), les mesures $\mu^{(r)}$ en sont également. Il a continué son étude en montrant des résultats d'indépendances statistiques : si on se donne b_1 et b_2 deux bases de numération entières premières entres-elles alors l'étude de $\Delta^{(r)}$ dans le premier système de numération est, en quelque sorte, indépendante de l'étude dans le second système.

Dans le cas de la représentation de Zeckendorf, des pistes ont commencé à apparaître dans la thèse [26] de Spiegelhofer qui, avec des arguments légèrement différents de ceux de Bésineau, a montré que la densité asymptotique de $\{n \in \mathbb{N} \ : \ \Delta^{(r)}(n) = d\}$ existe également. Cependant, il n'y a pas de formule semblable à (1) connue à ce jour.

En 2012, T.W. Cusick a proposé une conjecture qui a dynamisé la recherche dans ce domaine. C'est une conjecture qui concerne le système de numération binaire.

Conjecture (de Cusick.)**.** *Soit r un entier. Dans le cas binaire, $\mu^{(r)}(\mathbb{N}) > \frac{1}{2}$.*

Cette conjecture a de multiples traductions : elle revient à borner strictement par 2^r le nombre de racines du polynôme $(X+1)(X+2)\cdots(X+r)$ dans $\mathbb{Z}/2^{r+1}\mathbb{Z}$; elle se traduit également par des conjectures de divisibilités dans les colonnes ou les lignes de triangles de Pascal ; un autre énoncé équivalent peut faire intervenir des *écritures hyperbinaires* dont on ne parlera pas ici. Toutes ses traductions sont détaillées dans l'article [8] publié par Drmota, Kauers et Spiegelhofer.

Mais revenons-en à la conjecture énoncée plus haut : derrière la simplicité de son écriture se cache une réelle difficulté dans la recherche de sa démonstration. Cependant, avec un ordinateur et à l'aide de l'équation (1), on peut vérifier qu'elle est valable pour un grand nombre d'entiers : cela a été donc fait pour $r \leq 2^{30}$. Le dynamisme créé a amené à la découverte de bon nombre de résultats. Chronologiquement, Morgenbesser et Spiegelhofer ont démontré dans [19] une propriété étonnante valable en base entière : renverser l'ordre des chiffres de r ne modifie pas la loi $\mu^{(r)}$. Ensuite, Spiegelhofer a soutenu sa thèse [26] dans laquelle figurait le résultat d'existence de la densité asymptotique dans le cas Zeckendorf dont on a parlé plus haut mais également, à l'aide de son point de vue d'arithméticien, il a démontré que la Conjecture de Cusick est vraie pour un ensemble d'entiers de densité 1. Cette découverte a fait l'objet de l'article [8] co-écrit avec Drmota et Kauers. Ensuite, Emme a, dans le cadre de sa thèse, élargi la vision probabiliste de ce sujet. D'une collaboration avec Prikhod'ko est ressorti l'article [12] dans lequel est décrit le comportement asymptotique de la mesure $\mu^{(r)}$ dans le cas binaire à mesure que le nombre de motifs $\overline{01}$ augmente dans l'écriture binaire de r. En particulier, ils en déduisent un contrôle de la variance de $\mu^{(r)}$ selon ce nombre de motifs. Aussi, dans les articles [10] et [11], Emme et Hubert montrent que, dans le cas binaire, la mesure $\mu^{(r)}$ satisfait un Théorème Central Limite (TCL), à renormalisation près et à "un tirage presque sûr" de r. Leur technique consiste à montrer la convergence des moments de la loi renormalisée vers les moments de la loi normale centrée-réduite. Leur TCL implique que la valeur $\frac{1}{2}$ est un point d'accumulation pour la Conjecture de Cusick. Très récemment, en 2021, Spiegelhofer et Wallner ont prouvé dans [28] que la Conjecture de Cusick est vérifiée dès que l'écriture binaire de r fait intervenir assez de motifs $\overline{01}$ (au sens "plus qu'une quantité explicite").

Objectifs, approche et résultats.

Notre souhait de départ est de démontrer un TCL valable dans les systèmes de représentation de base entière et de Zeckendorf en explicitant davantage l'ensemble des r pour lesquels les mesures $\mu^{(r)}$ renormalisées s'approchent d'une loi normale. Les mesures étant des lois de probabilité, il paraît judicieux, pour les étudier, de se placer dans le cadre d'un espace probabilisé adapté au problème :

- en base entière b : les entiers b-adiques (ils généralisent les entiers naturels écrits en base b, mais avec une infinité de chiffres) ou
- en représentation de Zeckendorf : les entiers Z-adiques (ils généralisent de la même manière l'écriture des entiers naturel dans ce système de numération).

Ces espaces d'entiers b-adiques ou Z-adiques sont munis de transformations appelées "odomètres" qui modélisent l'addition de 1 dans le cadre de ces entiers généralisés. Ainsi, plutôt que de voir ce problème sur les entiers naturels, on le regarde plutôt par le prisme de l'espace probabilisé des entiers adiques. À ce niveau, il convient de distinguer les systèmes de numération.

En base entière $b \geq 2$

Il nous faut commencer par étendre $\Delta^{(r)}$ sur cet espace d'entiers b-adiques, ce qui est possible presque partout. Une autre étape consiste à prouver qu'en étendant à un espace plus grand que les entiers naturels, les lois de probabilités restent les mêmes. Nous construisons donc un pont solide entre le point de vue arithmétique et le point de vue dynamicien que nous avons choisi d'adopter. Un outil fondamental dans nos études est les tours de Rokhlin. Ces tours nous permettent d'obtenir des résultats de convergence ergodique et de (re)démontrer, par des moyens très différents de l'état de l'art, l'existence et la finitude des moments de tous ordres de $\Delta^{(r)}$ (Proposition 1.2.1 et Corollaire 1.2.2). Nous (re)prouvons également que $\Delta^{(r)}$ est une variable aléatoire centrée. Le lecteur pourra se référer à la Section 1.2 du premier chapitre pour de plus amples détails.

Ensuite, dans la Section 1.4, nous décomposons $\Delta^{(r)}$ en une somme de variables aléatoires formant un processus mélangeant. La décomposition se fait par "blocs".

Définition. *Un* bloc *dans l'écriture $\overline{r_\ell \cdots r_0}$ d'un entier $r \in \mathbb{N}$ est soit*

1. *une suite maximale de chiffres 0 consécutifs ou*
2. *une suite maximale de chiffres $b-1$ consécutifs ou*
3. *quand $b \geq 3$, un chiffre entre 1 et $b-2$.*

On définit, de plus, la quantité $\rho(r)$ comme étant le nombre de blocs dans l'écriture en base b de r.

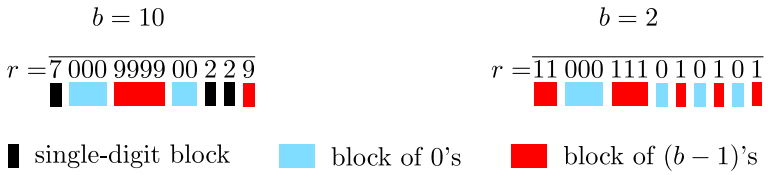


Figure 2: Deux exemples en base 2 et 10.

Les TCL dans le cadre de processus mélangeants ont été bien étudiés (voir par exemple [4, 7, 15, 16, 21, 22, 23, 24, 25, 29]). Néanmoins, nous avons besoin d'avoir un bon contrôle de la variance pour pouvoir obtenir un tel TCL. Le résultat suivant généralise le résultat en base 2 d'Emme et Priokhod'ko présent dans [12]. Il donne un contrôle de la variance de $\mu^{(r)}$ selon le nombre de blocs, valable en toute base.

Théorème (Theorem 1.1.2)**.** *Pour tout entier $r \geq 1$*

$$\frac{b}{4}\rho(r) \leq \mathrm{Var}(\mu^{(r)}) \leq 2b^2\rho(r).$$

Ce contrôle de variance a été obtenu dans la Section 1.3 et a pu être trouvé grâce à la relation de récurrence (1) sur $\mu^{(r)}$. En notant $\sigma_r := \sqrt{\mathrm{Var}(\mu^{(r)})} > 0$, il nous permet de considérer la loi renormalisée suivante

$$\forall d \in \mathbb{Z}, \qquad \widetilde{\mu}^{(r)}\left(\frac{d}{\sigma_r}\right) := \mu^{(r)}(d).$$

La mesure $\widetilde{\mu}^{(r)}$ est ainsi une mesure sur $\mathbb{R}$ concentrée sur les points de la forme $\frac{d}{\sigma_r}$ avec $d \in \mathbb{Z}$. Le contrôle de variance associé à de bonnes propriétés de mélange (Lemme 1.4.3) nous a permis d'établir le TCL avec vitesse suivant.

Théorème (Theorem 1.1.3,1.1.4)**.** *On a la convergence en loi*

$$\widetilde{\mu}^{(r)} \xrightarrow[\rho(r)\to+\infty]{d} \mathcal{N}(0,1).$$

Plus précisément, si F_r (respectivement F) est la fonction de répartition de $\widetilde{\mu}^{(r)}$ (respectivement $\mathcal{N}(0,1)$), alors on a l'existence d'une constante $K > 0$ telle que

$$\sup_{t\in\mathbb{R}} |F_r(t) - F(t)| \leq \frac{\widetilde{K}}{\rho(r)^{\frac{1}{8}}}.$$

Le lecteur pourra se référer aux Sections 1.1 et 1.5 du Chapitre 1 pour avoir des commentaires plus approfondis de ce résultat ainsi que sa démonstration. Nous précisons enfin que ce premier chapitre a fait l'objet d'une soumission à une revue.

En représentation de Zeckendorf

Le problème de variation de la somme des chiffres dans ce système de numération est plus difficile à appréhender qu'en base entière. Cette difficulté provient de l'étrangeté de l'addition dans ce système de numération : contrairement à l'addition en base entière, l'addition en représentation de Zeckendorf peut faire reculer des retenues ! Un travail de compréhension de ce phénomène est donc établi dans la Section 2.2. Ce comportement inhabituel de l'addition implique que certains résultats (quasiment) triviaux valables en base entière ne le sont plus dans ce système. Pire encore, ils peuvent devenir faux ! La formule (1) permettant, en base entière, le calcul de $\mu^{(r)}$ en est un exemple. Avant de pouvoir développer des résultats, il faut commencer par adapter la méthodologie introduite en base entière. Les tours correspondant à l'odomètre étant plus complexes qu'en base entière (elles sont deux à chaque étape), il faut adapter la méthodologie à ce nouveau système de numération. Ainsi, dans la Section 2.4 et les Sous-sections 2.3.1 et 2.2.2, on étend $\Delta^{(r)}$ aux entiers Z-adiques et on démontre que $\Delta^{(r)}$ est une variable aléatoire centrée possédant des moments finis de tous ordres.

De plus, ainsi qu'en base entière, on a décomposé $\Delta^{(r)}$ en une somme de variables aléatoires formant un processus mélangeant. Ce processus possède de bonnes propriétés de mélange que l'on décrit dans la Section 2.7, plus précisément, dans le Théorème 2.1.5. On peut alors espérer obtenir un TCL si nous arrivons à obtenir un contrôle de variance similaire à celui trouvé en base entière.

Cependant, un tel contrôle n'est, ici, pas clair du tout parce que des relations vérifiées par $\mu^{(r)}$ semblables à (1) ne sont (pour l'instant) pas connues. On peut néanmoins imaginer des pistes de réponses que l'on indiquera en Perspectives.

Malgré l'absence d'une formule semblable à (1), nous avons, grâce aux tours de Rokhkin, réussi à établir un algorithme effectuant le calcul de la mesure $\mu^{(r)}$. Une description de cet algorithme est donnée dans la Section 2.5 et, plus précisément, un code en pseudo-langage est écrit dans la Sous-section 2.5.2. Aussi, bien qu'étant énoncé et étudié dans ce système de représentation, cet algorithme s'adapte parfaitement au système en base entière. Il possède deux corollaires importants que nous énonçons. Le premier donne un contrôle de la queue de distribution de $\mu^{(r)}$.

Théorème (Théorème 2.1.1)**.** *Pour d assez petit dans $\mathbb{Z}$*

$$\mu^{(r)}(d-1) = \mu^{(r)}(d) \times \frac{1}{\varphi^2}.$$

On peut adapter ce résultat pour obtenir une formule similaire en base entière. L'autre corollaire est le résultat suivant.

Théorème (Théorème 2.1.2)**.** *Pour $\ell \geq 3$*

$$\mu^{(F_\ell)} = \mu^{(1)}.$$

Ce résultat est incontestablement trivial en base entière (en remplaçant F_ℓ par b^ℓ et via (1)) mais il ne l'est pas du tout en représentation de Zeckendorf à cause du comportement inhabituel des retenues que l'on a mentionné précédemment.

Chapitre 1

A central limit theorem for the variation of the sum of digits

Abstract. We prove a Central Limit Theorem for probability measures defined via the variation of the sum-of-digits function, in base $b \geq 2$. For $r \geq 0$ and $d \in \mathbb{Z}$, we consider $\mu^{(r)}(d)$ as the density of integers $n \in \mathbb{N}$ for which the sum of digits increases by d when we add r to n. We give a probabilistic interpretation of $\mu^{(r)}$ on the probability space given by the group of b-adic integers equipped with the normalized Haar measure. We split the base-b expansion of the integer r into so-called "blocks", and we consider the asymptotic behaviour of $\mu^{(r)}$ as the number of blocks goes to infinity. We show that, up to renormalization, $\mu^{(r)}$ converges to the standard normal law as the number of blocks of r grows to infinity. We provide an estimate of the speed of convergence. The proof relies, in particular, on a ϕ-mixing process defined on the b-adic integers.

Keywords. Sum of digits, Central Limit Theorem, b-adic odometer, ϕ-mixing.

MSC (2020). 11A63, 37A44 and 60F05.

1.1 Introduction

Throughout this article, *integers* mean elements of the set $\mathbb{N} := \{0, 1, 2, \cdots\}$, and b denotes a fixed integer, $b \geq 2$.

For an integer n, we consider the associated sequence of *digits* $(n_k) \in \{0, \cdots, b-1\}^{\mathbb{N}}$, finitely many of them being strictly positive, such that

$$n = \sum_{k \geq 0} n_k b^k.$$

For $n \neq 0$ and $\ell := \max\{k : n_k \neq 0\}$, we introduce the notation $\overline{n_\ell \cdots n_0} := n$, which generalises the usual way we write numbers in base 10, and which we refer to as the (*base* b) *expansion* of n. By convention, we set $\overline{0} := 0$. Then, we define the *sum-of-digits* function (in base b) as

$$s(n) := \sum_{k \geq 0} n_k.$$

A central object in our paper is the variation of the sum of digits when we add a fixed integer r to n: for $r, n \in \mathbb{N}$, we set

$$\Delta^{(r)}(n) := s(n+r) - s(n). \tag{1.1}$$

An interesting feature of $\Delta^{(r)}$ is that it gives the number of carries created during the addition $n + r$ in base b. To be more precise, if c is the number of carries then

$$\Delta^{(r)}(n) = s(r) - c(b-1). \tag{1.2}$$

Bésineau [3] proves that, for every $d \in \mathbb{Z}$, the following asymptotic density exists

$$\mu^{(r)}(d) := \lim_{N \to +\infty} \frac{1}{N} \left| \left\{ n < N : \Delta^{(r)}(n) = d \right\} \right|$$

and he studies these asymptotic densities through their correlation function. Actually, since $\sum_{d \in \mathbb{Z}} \mu^{(r)}(d) = 1$, the function $\mu^{(r)}$ defines a probability measure on $\mathbb{Z}$.

Morgenbesser and Spiegelhofer [19] show an amazing property: the measure $\mu^{(r)}$ remains the same if we reverse the order of the digits in the expansion of r. They call it the *reverse property*.

In the particular case $b = 2$, Emme and Prikhod'ko [12] show that the variance of $\mu^{(r)}$ is bounded from above and below by a constant multiplied by the number of blocks of 1's in the binary expansion of r. In Section 1.3, we extend this result to each $b \geq 2$.

Also in the binary case, Emme and Hubert [11] show that, for almost every sequence of integers $(r_n)_{n \in \mathbb{N}}$ written in binary and defined via a balanced Bernoulli process, the sequence of measures $(\mu^{(r_n)})_{n \in \mathbb{N}}$, after renormalization, converges in distribution to the standard normal law. The proof is done by computing all the moments of $\mu^{(r_n)}$ and by showing that, after renormalization, they converge to the moments of the standard normal law. In our paper, we prove a more accurate and more general Central Limit Theorem (CLT).

To do so, we study the variations of the sum of digits in the context of an appropriate probability space. We consider the compact additive group $(\mathbb{X}, +)$ of *b-adic integers*. The space $\mathbb{X}$ is endowed with the Borel σ-algebra and its normalized Haar measure $\mathbb{P}$.

We extend $\Delta^{(r)}$ almost everywhere on $\mathbb{X}$ and show in Section 1.2 (Proposition 1.2.1) that, for every $d \in \mathbb{Z}$

$$\mu^{(r)}(d) = \mathbb{P}\left(\{ x \in \mathbb{X} : \Delta^{(r)}(x) = d \} \right).$$

To state our main result, we have to define the notion of *blocks* in the base b expansion of an integer r.

Definition 1.1.1. *A* block *in the expansion $\overline{r_\ell \cdots r_0}$ of an integer $r \in \mathbb{N}$ is defined as follows: it is either*

1. *a maximal sequence of consecutive digits equal to 0 ("block of 0's") or*
2. *a maximal sequence of consecutive digits equal to $b-1$ ("block of $(b-1)$'s") or*
3. *when $b \geq 3$, a digit between 1 and $b-2$ ("single-digit block").*

We also define the quantity $\rho(r)$ as the number of blocks in the base b expansion of r.

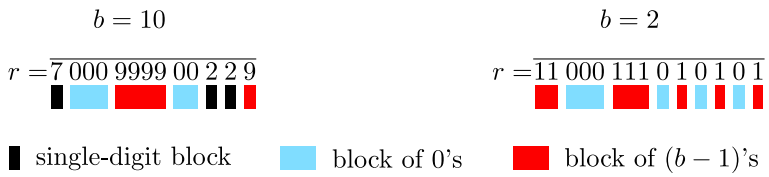


Figure 1.1: Examples of the decomposition in blocks in decimal and binary bases. On the left-hand side, $\rho(r) = 7$. On the right-hand side, $\rho(r) = 9$.

We specify that a block of 0's or of $(b-1)$'s of length 1 is not considered, in this paper, as a single-digit block.

The following theorem, which generalizes Emme and Prikhod'ko's result, states that the number of blocks $\rho(r)$ controls the variance of $\mu^{(r)}$.

Theorem 1.1.2. *For every integer $r \geq 1$*

$$\frac{b}{4}\rho(r) \leq \mathrm{Var}(\mu^{(r)}) \leq 2b^2\rho(r).$$

Now, we need to introduce, for an integer $r \geq 1$, the standard deviation $\sigma_r := \sqrt{\mathrm{Var}(\mu^{(r)})} > 0$ and the renormalized measure $\widetilde{\mu}^{(r)}$ which is the measure on $\mathbb{R}$ concentrated on the points of the form $\frac{d}{\sigma_r}$ $(d \in \mathbb{Z})$ and which satisfies

$$\forall d \in \mathbb{Z}, \qquad \widetilde{\mu}^{(r)}\left(\frac{d}{\sigma_r}\right) := \mu^{(r)}(d).$$

Our main result states that, for an integer $r \geq 1$, the renormalized measure $\widetilde{\mu}^{(r)}$ converges in distribution to the standard normal law as the number of blocks tends to infinity.

Theorem 1.1.3. *We have the convergence*

$$\widetilde{\mu}^{(r)} \xrightarrow[\rho(r)\to+\infty]{d} \mathcal{N}(0,1).$$

Theorem 1.1.3 can be seen as a direct consequence of the following theorem, which furthermore provides an estimation of the speed of convergence.

Theorem 1.1.4. *Let $h : \mathbb{R} \to \mathbb{R}$ be a thrice differentiable function with $||h'''||_\infty < \infty$. Let Z be a random variable following $\widetilde{\mu}^{(r)}$ and Y a standard normal random variable. Then*

$$\left|\mathbb{E}(h(Z)) - \mathbb{E}(h(Y))\right| = \underset{\rho(r)\to\infty}{O}\left(\frac{1}{\sqrt{\rho(r)}}\right). \tag{1.3}$$

Furthermore, if we denote by F_r (respectively F) the cumulative distribution function of $\widetilde{\mu}^{(r)}$ (respectively $\mathcal{N}(0,1)$), then there exists $\widetilde{K} > 0$ such that for every integer $r \geq 1$

$$\sup_{t\in\mathbb{R}} |F_r(t) - F(t)| \leq \frac{\widetilde{K}}{\rho(r)^{\frac{1}{8}}}. \tag{1.4}$$

A result in the same spirit has recently been published by Spiegelhofer and Wallner [28]. In the case of the base 2, they give a very accurate estimation of the measure $\mu^{(r)}(d)$ for every $d \in \mathbb{Z}$

$$\mu^{(r)}(d) = \frac{1}{\sigma_r\sqrt{2\pi}} e^{-\frac{d^2}{2\sigma_r^2}} + \underset{\rho(r)\to\infty}{O}\left(\rho(r)^{-1}(\log(\rho(r))^4\right). \tag{1.5}$$

Their result is proved using a combination of several techniques such as recurrence relations, cumulant generating functions, and integral representations. It seems possible but extremely difficult to generalize (1.5) to other bases. It also implies a CLT when $\rho(r)$ tends to infinity: using (1.5), it is possible to show that for every real numbers $a < b$

$$\widetilde{\mu}^{(r)}([a,b]) \xrightarrow[\rho(r)\to\infty]{} \frac{1}{\sqrt{2\pi}} \int_a^b e^{\frac{-t^2}{2}} \mathrm{d}t$$

with a speed of convergence of $\frac{\log^4(\rho(r))}{\rho(r)^{\frac{1}{2}}}$. However, it is not clear how we can get a speed of convergence of $F_r(t)$ to $F(t)$. On our side, we use a drastically different approach which applies directly in any base, relying on the concept of *ϕ-mixing* process and on a result from Sunklodas [29].

Roadmap

Section 1.2 is devoted to placing the study of the measures $\mu^{(r)}$ in the context of the odometer on the set $\mathbb{X}$ of b-adic integers. We extend $\Delta^{(r)}$ almost everywhere on $\mathbb{X}$ and we show that the convergence

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n<N} f(\Delta^{(r)}(n)) = \int_{\mathbb{X}} f(\Delta^{(r)}(x))\mathrm{d}\mathbb{P}(x)$$

(where $\mathbb{P}$ is the normalized Haar measure on $\mathbb{X}$) is satisfied for functions $f : \mathbb{Z} \to \mathbb{C}$ of polynomial growth (Proposition 1.2.1) and, more generally, for functions f such that $f \circ \Delta^{(r)}$ is integrable (Proposition 1.2.4). We deduce from Proposition 1.2.1 that $\mu^{(r)}(d) = \mathbb{P}(\{x \in \mathbb{X} : \Delta^{(r)}(x) = d\})$ and that $\mu^{(r)}$ has finite moments. In particular, we show that $\mu^{(r)}$ is of zero-mean.

In Section 1.3, we focus on the second moment of $\mu^{(r)}$. First, we establish in Proposition 1.3.1 an inductive relation between the measures in the spirit of Bésineau's result [3, p.13]. From that, we deduce an inductive relation on the variance of the measures (Lemma 1.3.3). Then, we prove the estimation of the variance stated in Theorem 1.1.2.

In Section 1.4, we build a finite sequence of random variables associated to the addition of some integer r that will be used to prove Theorem 1.1.4. We estimate the ϕ-mixing coefficients for this sequence.

The last section is devoted to the proof of Theorem 1.1.4. We show how we can apply a result from Sunklodas [29, Theorem 1] giving a speed of convergence in the CLT for ϕ-mixing process.

Acknowledgment

We thank M. El Machkouri for fruithful discussions and references about the speed of convergence in CLT for mixing sequences.

1.2 Odometer and sum-of-digits function

1.2.1 Unique ergodicity of the b-adic odometer

We define $\mathbb{X}$ as the space of $b-$*adic integers*, that is the space $\{0,\cdots,b-1\}^{\mathbb{N}}$. Coordinates of a b-adic integer $x\in\mathbb{X}$ are interpreted as digits in base b: elements of $\mathbb{X}$ can be viewed as "generalized integers having possibly infinitely many non zero digits in base b". To comply with the usual writing of numbers in base 10, an element $x=(x_n)_{n\in\mathbb{N}}\in\mathbb{X}$ will be represented as a left-infinite sequence $(\cdots,x_1,x_0)$, x_0 being the units digit. The space $\mathbb{X}$ is compact for the product topology. The set of integers $\mathbb{N}$ can be identified with the subset of sequences with finite support. More precisely, using the inclusion function

$$i:\overline{n_\ell\cdots n_0}\in\mathbb{N}\longmapsto(\cdots,0,n_\ell,\cdots,n_1,n_0)\in\mathbb{X}$$

we identify $\mathbb{N}$ and $i(\mathbb{N})$.

$\mathbb{X}$ is equipped with an addition which extends the usual addition on $\mathbb{N}$ and turns $\mathbb{X}$ into an Abelian group. For $x=(\cdots,x_0)$ and $y=(\cdots,y_0)$ in $\mathbb{X}$, $(x+y)$ is determined recursively by the following process, in which we generate a sequence of carries $(c_\ell)_{\ell\geq0}\in\{0,1\}^{\mathbb{N}}$.

- Initialisation:
  if $x_0+y_0<b$ then we set $(x+y)_0:=x_0+y_0$ and $c_0:=0$,
  else we set $(x+y)_0:=x_0+y_0-b$ and $c_0:=1$.

- Induction step: once, for some $\ell\geq1$, we have computed $(x+y)_i$ and c_i for $i=0,\cdots,\ell-1$,
  if $x_\ell+y_\ell+c_{\ell-1}<b$ then we set $(x+y)_\ell:=x_\ell+y_\ell+c_{\ell-1}$ and $c_\ell:=0$,
  else we set $(x+y)_\ell:=x_\ell+y_\ell+c_{\ell-1}-b$ and $c_\ell:=1$.

Now, since 1 belongs to $\mathbb{N}\subset\mathbb{X}$, we can consider the application $T:\mathbb{X}\to\mathbb{X}$

$$T:x\longmapsto T(x):=x+1,$$

which is usually refered to as the b-*adic odometer*. It is well-known that T is a homeomorphism and so $(\mathbb{X},T)$ is a topological dynamical system [13]. For $\ell\geq0$ and for integers $r_\ell,\cdots,r_0\in\{0,\cdots,b-1\}$, we define the *cylinder* $C_{r_\ell\cdots r_0}$ as the set of sequences x such that $x_i=r_i$ for $i=0,\cdots,\ell$. We observe that the image by T of a cylinder is another cylinder : for integers $r_\ell,\cdots,r_0\in\{0,\cdots,b-1\}$, if there exists a minimal index $i\in\{0,\cdots,\ell\}$ such that $r_i\neq b-1$ then $TC_{r_\ell\cdots r_0}=C_{r_\ell\cdots r_{i+1}(1+r_i)0^i}$, otherwise $TC_{r_\ell\cdots r_0}=C_{0^{\ell+1}}$. Also, we define the *Rokhlin tower of order* $\ell\geq0$ as the family

$$\left(C_{0^{\ell+1}},TC_{0^{\ell+1}},\cdots,T^{b^{\ell+1}-1}C_{0^{\ell+1}}\right)$$

where $(T^jC_{0^{\ell+1}})_{0\leq j\leq b^{\ell+1}-1}$ form a partition of $\mathbb{X}$. We commonly represent this family as a tower as shown in Figure 1.2.

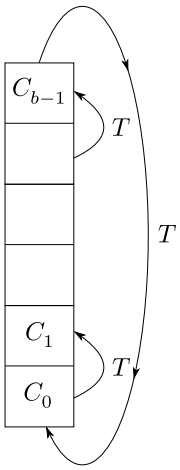


Figure 1.2: Behavior of T on the Rokhlin tower of order 0.

It is classical that the sequence of towers can be constructed with the so-called *Cut-and-Stack* inductive process, as illustrated in Figure 1.3.

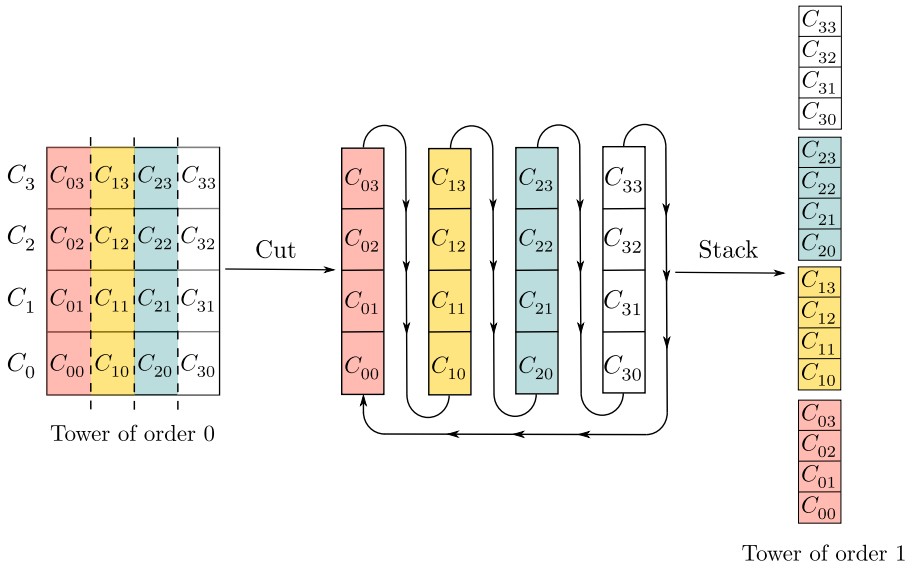


Figure 1.3: How to construct the tower of order 1 of T from the tower of order 0 in base $b = 4$.

By looking at the behavior of T on these towers, one can show that if $\mathbb{P}$ is a $T-$invariant probability measure on $(\mathbb{X}, T)$, then $\mathbb{P}$ gives the same measure to each level in a given tower: for every $\ell \geq 0$ and $r_0, \cdots, r_\ell \in \{0, \cdots, b-1\}$

$$\mathbb{P}\left(C_{r_\ell \ldots r_0}\right) = \frac{1}{b^{\ell+1}}.$$

Since the cylinders generate the Borel σ-algebra, $\mathbb{P}$ is uniquely determined by these values on the cylinders, hence $(\mathbb{X}, T)$ is a uniquely ergodic dynamical system. We observe that choosing x in $\mathbb{X}$ according to the unique T-invariant law $\mathbb{P}$ means choosing its digits independently according to the uniform law on $\{0, \cdots, b-1\}$. We also note that $\mathbb{P}$ is the normalized Haar measure on $\mathbb{X}$.

For x in $\mathbb{X}$, we define the sequence of empirical probability measures along the (beginning of the) orbit of x: for every $N \geq 1$, we set

$$\epsilon_N(x) := \frac{1}{N} \sum_{0 \leq n < N} \delta_{T^n x}$$

(where δ_y denotes the Dirac measure on $y \in \mathbb{X}$).
Since the space of probability measures on $\mathbb{X}$ is compact for the weak-$*$ topology, every subsequential limit of $(\epsilon_N(x))$ is a T-invariant probability measure. By the uniqueness of the T-invariant probability measure, for every $x \in \mathbb{X}$ we have $\epsilon_N(x) \to \mathbb{P}$. In other words, we have the convergence

$$\forall x \in \mathbb{X}, \ \ \forall f \in \mathcal{C}(\mathbb{X}), \qquad \frac{1}{N} \sum_{0 \leq n < N} f(T^n x) \xrightarrow[N \to +\infty]{} \int_{\mathbb{X}} f \mathrm{d}\mathbb{P}. \tag{1.6}$$

We will be interested here in the special case $x = 0$ because $\mathbb{N} = \{T^n 0 : n \in \mathbb{N}\}$. Then (1.6) becomes

$$\forall f \in \mathcal{C}(\mathbb{X}), \qquad \frac{1}{N} \sum_{0 \leq n < N} f(n) \xrightarrow[N \to +\infty]{} \int_{\mathbb{X}} f \mathrm{d}\mathbb{P}. \tag{1.7}$$

Equation (1.7) shows that, for a continuous function f, averaging f over $\mathbb{N}$ (for the natural density) amounts to averaging over $\mathbb{X}$ (for $\mathbb{P}$). The next section shows how this convergence can be extended to some non-continuous functions related to the sum-of-digits function.

1.2.2 Sum of digits on the odometer

For every integer k, we define $s_k : \mathbb{X} \to \mathbb{Z}$ as the sum of the first $(k+1)$ digits function, that is to say

$$s_k(x) := x_0 + \cdots + x_k.$$

Let $r \in \mathbb{N}$. We define the functions $\Delta_k^{(r)} : \mathbb{X} \to \mathbb{Z}$ by

$$\Delta_k^{(r)}(x) := s_k(x + r) - s_k(x).$$

The functions $\Delta_k^{(r)}$ are well-defined, continuous (and bounded) on $\mathbb{X}$. By (1.7), we have

$$\frac{1}{N} \sum_{n < N} \Delta_k^{(r)}(n) = \frac{1}{N} \sum_{n < N} \Delta_k^{(r)}(T^n 0) \xrightarrow[N \to +\infty]{} \int_{\mathbb{X}} \Delta_k^{(r)} \mathrm{d}\mathbb{P}. \tag{1.8}$$

Although the sum-of-digits function s is not well defined on $\mathbb{X}$, we can extend the function $\Delta^{(r)}$ defined in (1.1) on the set of $x \in \mathbb{X}$ for which the number of different digits between x and $x + r$ is finite. This subset contains the b-adic integers x such that there exists an index $k \geq \max(\{\ell : r_\ell \neq 0\}$ such that $x_k \neq b - 1$. So, except for a finite number of b-adic integers, we can define

$$\Delta^{(r)}(x) := \lim_{k \to \infty} \Delta_k^{(r)}(x).$$

Remark 1. *Let $t \geq u$ be two integers. For every integer k we have the decomposition formula*

$$\Delta_k^{(t+u)} = \Delta_k^{(t)} + \Delta_k^{(u)} \circ T^t. \tag{1.9}$$

So, taking $\mathbb{P}$-almost everywhere the limit when k tends to infinity, we get

$$\Delta^{(t+u)} = \Delta^{(t)} + \Delta^{(u)} \circ T^t \ \ (\mathbb{P}\text{-}a\text{-}s.). \tag{1.10}$$

Then, by induction on t, we deduce

$$\Delta^{(t)} = \Delta^{(1)} + \Delta^{(1)} \circ T + \cdots + \Delta^{(1)} \circ T^{t-1} \ \ (\mathbb{P}\text{-}a\text{-}s.). \tag{1.11}$$

We observe that $\Delta^{(r)}$ also satisfies (1.2): for each $x \in \mathbb{X}$ for which $\Delta^{(r)}(x)$ is well defined, we have

$$\Delta^{(r)}(x) = s(r) - c(b-1), \tag{1.12}$$

where $c := \sum_{\ell \geq 0} c_\ell < \infty$ is the total number of carries generated during the computation of $x + r$. Unfortunately, $\Delta^{(r)}$ is not continuous like the functions $\Delta_k^{(r)}$, it is not even bounded on $\mathbb{X}$, but we have the following result about functions with polynomial growth.

Proposition 1.2.1. *Let $r \geq 1$ and $f : \mathbb{Z} \to \mathbb{C}$. Assume that there exist $\alpha \geq 1$ and C in $\mathbb{R}_+^*$ such that for every $n \in \mathbb{Z}$*

$$|f(n)| \leq C|n|^\alpha + |f(0)|. \tag{1.13}$$

Then $f \circ \Delta^{(r)} \in L^1(\mathbb{P})$ and we have the convergence

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N} \sum_{n<N} f(\Delta^{(r)}(n)) &= \int_{\mathbb{X}} f(\Delta^{(r)}(x)) \mathrm{d}\mathbb{P}(x) \\ &= \lim_{k\to\infty} \int_{\mathbb{X}} f(\Delta_k^{(r)}(x)) \mathrm{d}\mathbb{P}(x). \end{aligned}$$

Corollary 1.2.2. *For every $d \in \mathbb{Z}$*

$$\begin{aligned} \mu^{(r)}(d) &:= \lim_{N\to\infty} \frac{1}{N} \left| \left\{ n < N : \Delta^{(r)}(n) = d \right\} \right| \\ &= \mathbb{P}\left(\left\{ x \in \mathbb{X} : \Delta^{(r)}(x) = d \right\} \right). \end{aligned} \tag{1.14}$$

Moreover, $\Delta^{(r)}$ has zero-mean and has finite moments.

In particular, we recover Bésineau's result on the existence of the asymptotic density and the fact that $\sum_{d\in\mathbb{Z}} \mu^{(r)}(d) = 1$.

Remark 2. *Using trivial arguments, Proposition 1.2.1 and Corollary 1.2.2 are also true when $r = 0$. We observe that $\mu^{(0)} = \delta_0$.*

Before proving this proposition and its corollary, we need the following lemma.

Lemma 1.2.3. *Let $r \geq 1$. For $N \in \mathbb{N}^*$, for $k \in \mathbb{N}$ and $d, d' \in \mathbb{Z}$, we have the inequality*

$$\begin{aligned} \frac{1}{N}\left|\{n < N : (\Delta^{(r)}(n), \Delta_k^{(r)}(n)) = (d, d')\}\right| \\ \leq rb\,\mathbb{P}\Big(\{x \in \mathbb{X} : (\Delta^{(r)}(x), \Delta_k^{(r)}(x)) = (d, d')\}\Big). \end{aligned} \tag{1.15}$$

In particular, we have

$$\frac{1}{N}\left|\{n < N : \Delta^{(r)}(n) = d\}\right| \leq rb\,\mathbb{P}\Big(\{x \in \mathbb{X} : \Delta^{(r)}(x) = d\}\Big). \tag{1.16}$$

Proof. Of course, (1.15) implies (1.16) so we just need to prove (1.15). We fix $k \in \mathbb{N}$. For every $\ell \in \mathbb{N}$, let V_ℓ be the set of the values reached by the couple $(\Delta^{(r)}, \Delta_k^{(r)})$ on the first $b^{\ell+1} - r$ levels of the Rokhlin tower of order ℓ (see Figure 1.4). Of course, if $b^{\ell+1} - r \leq 0$ then $V_\ell := \emptyset$. Otherwise, we observe that V_ℓ is a finite set. Indeed, the first $b^{\ell+1} - r$ levels correspond to the b-adic integers x such that, when we add r, the carry propagation does not go beyond the first $\ell + 1$ digits. Since these digits are fixed on a level of the Rokhlin tower of order ℓ, except for the last r levels, $\Delta^{(r)}$ and $\Delta_k^{(r)}$ are constant on each such level. We observe that the sequence $(V_\ell)_{\ell \geq 0}$ is increasing for the inclusion.

Now, for $d, d' \in \mathbb{Z}$, there are 2 cases.

1. If $(d, d') \notin \cup_{\ell \geq 0} V_\ell$, then for each $n \in \mathbb{N}$ we have $(\Delta^{(r)}(n), \Delta_k^{(r)}(n)) \neq (d, d')$. Indeed, for each $n \in \mathbb{N}$, there exists a smallest integer ℓ such that n is in the first $b^{\ell+1} - r$ levels of the tower of order ℓ, hence $(\Delta^{(r)}(n), \Delta_k^{(r)}(n))$
$\in V_\ell$. In this case, (1.15) is trivial.

2. If $(d, d') \in \cup_{\ell \geq 0} V_\ell$ then there exists a unique $\ell \geq 0$ such that $(d, d') \in V_\ell \backslash V_{\ell-1}$ (with the convention $V_{-1} := \emptyset$). Since $(\Delta^{(r)}, \Delta_k^{(r)})$ is constant on each of the first $b^{\ell+1} - r$ levels of the tower, it takes the value (d, d') on at least one whole such level of measure $\frac{1}{b^{\ell+1}}$. So, we have

$$\mathbb{P}\Big(\{x \in \mathbb{X} : (\Delta^{(r)}(x), \Delta_k^{(r)}(x)) = (d, d')\}\Big) \geq \frac{1}{b^{\ell+1}}.$$

   Also, since the couple (d, d') does not appear in the first levels of the previous tower ($(d, d') \notin V_{\ell-1}$), we claim that, for every $N \geq 1$

$$\frac{1}{N}\left|\{n < N : (\Delta^{(r)}(n), \Delta_k^{(r)}(n)) = (d, d')\}\right| \leq \frac{r}{b^\ell}.$$

   Indeed,

   (a) If $r \geq b^\ell$, then the inequality is then trivial.

   (b) If $r < b^\ell$ then, since (d, d') is not in $V_{\ell-1}$, (d, d') can only appear inside the r highest levels of the tower of order $\ell - 1$. So, if we note C the union of these r highest levels, using the fact that 0 is in the first level of the tower, we have

$$\begin{aligned} \frac{1}{N}\left|\{0 \leq n < N : (\Delta^{(r)}(n), \Delta_k^{(r)}(n)) = (d, d')\}\right| \\ \leq \frac{1}{N}\left|\{0 \leq n < N : T^n 0 \in C\}\right| \\ \leq \frac{r}{b^\ell}. \end{aligned}$$

Combining both inequalities gives (1.15).

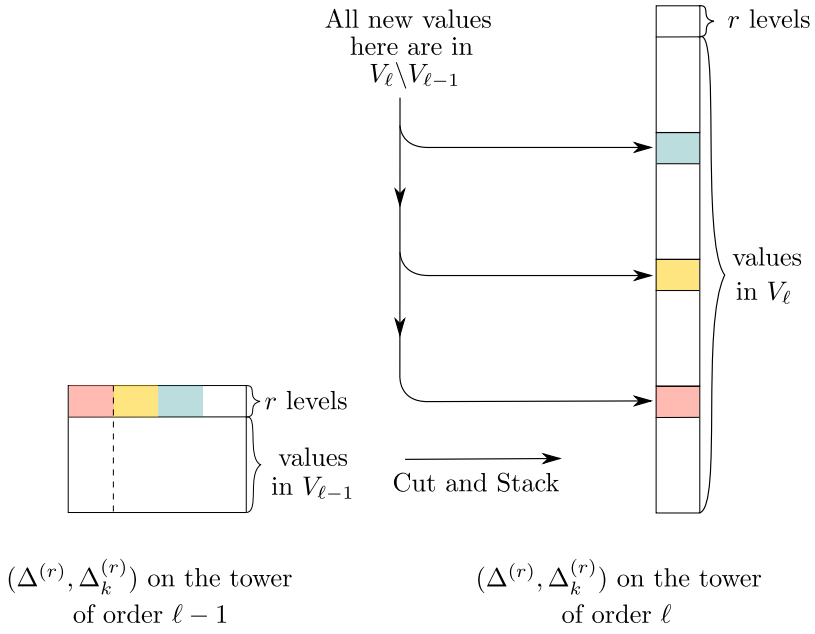


Figure 1.4: Visual description of $V_{\ell-1}$, V_ℓ and $V_\ell \setminus V_{\ell-1}$.

□

Proof of Proposition 1.2.1. Let $\varepsilon > 0$. For any integer k, we have

$$\left| \frac{1}{N} \sum_{n<N} f(\Delta^{(r)}(n)) - \int_{\mathbb{X}} f(\Delta^{(r)}(x)) \mathrm{d}\mathbb{P}(x) \right| \leq A_1 + A_2 + A_3,$$

where

$$A_1 := \frac{1}{N} \sum_{n<N} \left| f(\Delta^{(r)}(n)) - f(\Delta_k^{(r)}(n)) \right|,$$
$$A_2 := \left| \frac{1}{N} \sum_{n<N} f(\Delta_k^{(r)}(n)) - \int_{\mathbb{X}} f(\Delta_k^{(r)}(x)) \mathrm{d}\mathbb{P}(x) \right|,$$
$$A_3 := \int_{\mathbb{X}} \left| f(\Delta^{(r)}(x)) - f(\Delta_k^{(r)}(x)) \right| \mathrm{d}\mathbb{P}(x).$$

For A_1, using Lemma 1.2.3, we get

$$\begin{aligned}
A_1 &= \frac{1}{N}\sum_{n<N}\sum_{j,j'\in\mathbb{Z}} \Big|f(j)-f(j')\Big| \mathbb{1}_{(j,j')}\left(\Delta^{(r)}(n),\Delta_k^{(r)}(n)\right) \\
&= \sum_{j,j'\in\mathbb{Z}} \Big|f(j)-f(j')\Big| \frac{1}{N}\sum_{n<N} \mathbb{1}_{(j,j')}\left(\Delta^{(r)}(n),\Delta_k^{(r)}(n)\right) \\
&\leq rb \sum_{j,j'\in\mathbb{Z}} \Big|f(j)-f(j')\Big| \,\mathbb{P}\Big(\Delta^{(r)}(n)=j,\Delta_k^{(r)}(n)=j'\Big) \\
&= rb\int_{\mathbb{X}} \Big|f(\Delta^{(r)}(x))-f(\Delta_k^{(r)}(x))\Big| \mathrm{d}\mathbb{P}(x) = rbA_3.
\end{aligned}$$

Hence, controlling A_3 also enables us to control A_1. For this, we note that the integrand in A_3 converges $\mathbb{P}$-almost everywhere to 0, therefore it is enough to show that the dominated convergence theorem applies. To find a good dominant function, we write using (1.13)

$$\Big|f\circ\Delta_k^{(r)}(x)\Big| \leq C\left|\Delta_k^{(r)}(x)\right|^{\alpha} + \Big|f(0)\Big|.$$

We get from (1.9)

$$\Delta_k^{(r)} = \Delta_k^{(1)} + \Delta_k^{(1)}\circ T + \cdots + \Delta_k^{(1)}\circ T^{r-1},$$

then we use the multinomial theorem (we can suppose $\alpha\in\mathbb{N}$) to write

$$\left|\Delta_k^{(1)}(x)+\cdots+\Delta_k^{(1)}\circ T^{r-1}(x)\right|^{\alpha} = \sum_{j_0+\ldots+j_{r-1}=\alpha}\binom{\alpha}{j_0,\ldots,j_{r-1}}\prod_{i=0}^{r-1}\left|\Delta_k^{(1)}\circ T^i(x)\right|^{j_i}.$$

We now use Young's inequality to get

$$\prod_{i=0}^{r-1}\left|\Delta_k^{(1)}\circ T^i(x)\right|^{j_i} \leq \frac{1}{r}\sum_{i=0}^{r-1}\left|\Delta_k^{(1)}\circ T^i(x)\right|^{rj_i}.$$

Then, we define, for $i=0,\cdots,r-1$, $g_i(x) := \sup_{k\in\mathbb{N}}|\Delta_k^{(1)}\circ T^i(x)|$ in order to get the inequality

$$\Big|f\circ\Delta_k^{(r)}(x)\Big| \leq C\sum_{j_0+\cdots+j_{r-1}=\alpha}\sum_{i=0}^{r-1}\frac{\binom{\alpha}{j_0,\cdots,j_{r-1}}}{r}g_i(x)^{rj_i} + \Big|f(0)\Big|.$$

Moreover, when it is well defined, we have $\Delta^{(r)}\circ T^i = \lim_{k\to\infty}\Delta_k^{(r)}\circ T^i$ therefore, we also get $\left|\Delta^{(r)}\circ T^i\right| \leq g_i$ which yields the similar inequality

$$\Big|f\circ\Delta^{(r)}(x)\Big| \leq C\sum_{j_0+\cdots+j_{r-1}=\alpha}\sum_{i=0}^{r-1}\frac{\binom{\alpha}{j_0,\cdots,j_{r-1}}}{r}g_i(x)^{rj_i} + \Big|f(0)\Big|.$$

We just need to show that $g_i^{rj_i}$ is integrable for the measure $\mathbb{P}$. It is equivalent to show that $\sum_m \mathbb{P}\left(\{x \in \mathbb{X} : g_i(x)^{rj_i} > m\}\right)$ is a convergent series. We have

$$\begin{aligned} g_i(x) > m^{\frac{1}{rj_i}} &\Leftrightarrow \sup_{k\in\mathbb{N}} |\Delta_k^{(1)} \circ T^i(x)| > m^{\frac{1}{rj_i}} \\ &\Leftrightarrow \exists k \in \mathbb{N},\ \ |\Delta_k^{(1)}(T^i x)| > m^{\frac{1}{rj_i}}. \end{aligned}$$

From (1.12), this condition is true if and only if the addition $T^i x + 1$ creates sufficently many carries: strictly more than $\lfloor \frac{m^{\frac{1}{rj_i}}+1}{b-1} \rfloor$. But, we know that when we add 1 to $T^i x$, the number of carries created by the addition is the number of $(b-1)$'s at the right-hand side of the expansion of $T^i x$. So, because of the T-invariance of $\mathbb{P}$

$$\mathbb{P}\left(\{x \in \mathbb{X} : g_i(x) > m^{\frac{1}{rj_i}}\}\right) \leq \left(\frac{1}{b}\right)^{\left\lfloor \frac{m^{\frac{1}{rj_i}}+1}{b-1} \right\rfloor}.$$

The right quantity is the general term of a convergent series which shows that $g_i^{rj_i} \in L^1(\mathbb{P})$. So the dominated convergence theorem can be applied and, for k large enough, $A_1 + A_3 \leq \frac{\varepsilon}{2}$ for every $N \geq 1$.
Now, once we have fixed such a k, for N large enough, A_2 is bounded by $\frac{\varepsilon}{2}$ because of (1.7) and the continuity of $\Delta_k^{(r)}$ and of f. The convergence in the statement is thus proved.

Note that the argument of the dominated convergence theorem also proves that $f \circ \Delta^{(r)} \in L^1(\mathbb{P})$ and $\int_{\mathbb{X}} f \circ \Delta^{(r)} \mathrm{d}\mathbb{P} = \lim_{k\to\infty} \int_{\mathbb{X}} f \circ \Delta_k^{(r)} \mathrm{d}\mathbb{P}$. □

Proof of Corollary 1.2.2. We just apply Proposition 1.2.1 with particular functions f. First, for $d \in \mathbb{Z}$, we use the function $f = \mathbb{1}_{\{d\}}$. It gives

$$\mu^{(r)}(d) = \mathbb{P}\left(\left\{x \in \mathbb{X} : \Delta^{(r)}(x) = d\right\}\right).$$

Then, we take f as the identity function on $\mathbb{Z}$ for which (1.13) is clearly satisfied. Also, for every integer k, we have by T-invariance of $\mathbb{P}$

$$\int_{\mathbb{X}} \Delta_k^{(r)} \mathrm{d}\mathbb{P} = \int_{\mathbb{X}} s_k \circ T^r \mathrm{d}\mathbb{P} - \int_{\mathbb{X}} s_k \mathrm{d}\mathbb{P} = 0.$$

We then deduce that

$$\sum_{d\in\mathbb{Z}} d\,\mu^{(r)}(d) = \int_{\mathbb{X}} \Delta^{(r)} \mathrm{d}\mathbb{P} = \lim_{k\to\infty} \int_{\mathbb{X}} \Delta_k^{(r)} \mathrm{d}\mathbb{P} = 0. \tag{1.17}$$

Finally, we use, for every $j \geq 2$, the function $f(n) = n^j$ which satisfies (1.13). This gives the existence of moments of order j for $\Delta^{(r)}$. □

More generally, we have the following convergence.

Proposition 1.2.4. *Let $r \geq 1$ and $f : \mathbb{Z} \to \mathbb{C}$ be such that $f \circ \Delta^{(r)} \in L^1(\mathbb{P})$. Then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n<N} f(\Delta^{(r)}(n)) = \int_{\mathbb{X}} f(\Delta^{(r)}(x)) \mathrm{d}\mathbb{P}(x).$$

Proof of Proposition 1.2.4. By (1.12), the values reached by $\Delta^{(r)}$ are of the form $a_k^{(r)} := s(r) - k(b-1)$, $k \geq 0$, and we have

$$\int_{\mathbb{X}} f\left(\Delta^{(r)}(x)\right) \mathrm{d}\mathbb{P}(x) = \sum_{k\geq 0} f\left(a_k^{(r)}\right) \mathbb{P}\left(\{x \in \mathbb{X} : \Delta^{(r)}(x) = a_k^{(r)}\}\right). \tag{1.18}$$

On the other hand, we write

$$\begin{aligned}\frac{1}{N} \sum_{n<N} f\left(\Delta^{(r)}(n)\right) &= \frac{1}{N} \sum_{n<N} f\left(\Delta^{(r)}(n)\right) \sum_{k\geq 0} \mathbb{1}_{\left\{a_k^{(r)}\right\}}\left(\Delta^{(r)}(n)\right) \\ &= \sum_{k\geq 0} \underbrace{f\left(a_k^{(r)}\right) \frac{1}{N} \sum_{n<N} \mathbb{1}_{\left\{a_k^{(r)}\right\}}\left(\Delta^{(r)}(n)\right)}_{:=u_{N,k}}.\end{aligned}$$

We conclude by applying the dominated convergence theorem in the metric space $\ell^1(\mathbb{N})$ endowed with the counting measure to show

$$\sum_{k\geq 0} u_{N,k} \xrightarrow[N\to+\infty]{} \sum_{k\geq 0} f\left(a_k^{(r)}\right) \mathbb{P}\left(\{x \in \mathbb{X} : \Delta^{(r)}(x) = a_k^{(r)}\}\right).$$

1. The pointwise limit on $\mathbb{N}$ of $u_{N,k}$ is $f\left(a_k^{(r)}\right) \mathbb{P}\left(\{x \in \mathbb{X} : \Delta^{(r)}(x) = a_k^{(r)}\}\right)$ by (1.14).
2. A dominant function is given using Lemma 1.2.3: for all k and N we have

$$u_{N,k} \leq g(k) := \left|f\left(a_k^{(r)}\right)\right| rb\, \mathbb{P}\Big(\{x \in \mathbb{X} : \Delta^{(r)}(x) = a_k^{(r)}\}\Big).$$

   By (1.18) and the hypothesis that $f \circ \Delta^{(r)}$ is integrable, $\sum g(k)$ is convergent.

□

We have shown that, for any integer r, $\Delta^{(r)}$ has finite moments. The next section will focus on the second moment, which is the one we are most interested in for our CLT.

1.3 Variance of $\mu^{(r)}$

1.3.1 Inductive relation on the measures

Given $r \in \mathbb{N}$, there exist $\widetilde{r} \in \mathbb{N}$ and $0 \leq r_0 \leq b-1$ such that $r = b\widetilde{r} + r_0$. The integer r_0 is actually the units digit of the expansion of r and, if $r \geq b$, $\widetilde{r}$ corresponds to the integer whose expansion is obtained by erasing r_0 in the expansion of r. The expansion of $\widetilde{r}$ is then one digit shorter than the expansion of r. If $r < b$ then, of course, $r_0 = r$ and $\widetilde{r} = 0$. First of all, we have a well known inductive relation on the length of the expansion of r.

Proposition 1.3.1. *For $\widetilde{r} \in \mathbb{N}$, $0 \leq r_0 \leq b-1$ and $d \in \mathbb{Z}$*

$$\mu^{(b\widetilde{r}+r_0)}(d) = \frac{b-r_0}{b}\mu^{(\widetilde{r})}(d-r_0) + \frac{r_0}{b}\mu^{(\widetilde{r}+1)}(d+b-r_0). \tag{1.19}$$

Proof. Let $x = (\cdots, x_1, x_0) \in \mathbb{X}$. We define $\widetilde{x} := (\cdots, x_1)$. Let us consider the computation of the digits of $x + r$, where $r := b\widetilde{r} + r_0 = \overline{r_\ell \cdots r_0}$

$$\begin{array}{cccccc} & \cdots & x_\ell & \cdots & x_1 & x_0 \\ + & & r_\ell & \cdots & r_1 & r_0 \\ \hline = & & & \cdots & & (x+r)_0 \end{array}$$

If $x_0 + r_0 < b$, no carry is created and $\Delta^{(b\widetilde{r}+r_0)}(x) = r_0 + \Delta^{(\widetilde{r})}(\widetilde{x})$. Otherwise, $x_0 + r_0 \geq b$ and we have to subtract b from the units digit of the result: we are left with the addition of $\widetilde{x}$ and $\widetilde{r}+1$: so $\Delta^{(b\widetilde{r}+r_0)}(x) = r_0 - b + \Delta^{(\widetilde{r}+1)}(\widetilde{x})$. To sum up, we have

$$\Delta^{(b\widetilde{r}+r_0)}(x) = \begin{cases} r_0 + \Delta^{(\widetilde{r})}(\widetilde{x}) & \text{if } x_0 + r_0 < b, \\ r_0 - b + \Delta^{(\widetilde{r}+1)}(\widetilde{x}) & \text{otherwise.} \end{cases}$$

Now, let $d \in \mathbb{Z}$. We partition the set $\{x \in \mathbb{X} : \Delta^{(b\widetilde{r}+r_0)}(x) = d\}$ according to the value of x_0

$$\begin{aligned} \{x \in \mathbb{X} : \Delta^{(b\widetilde{r}+r_0)}(x) = d\} = & \bigcup_{j=0}^{b-r_0-1} \{x \in \mathbb{X} : x_0 = j \text{ and } \Delta^{(\widetilde{r})}(\widetilde{x}) = d - r_0\} \\ & \bigcup \bigcup_{j=b-r_0}^{b-1} \{x \in \mathbb{X} : x_0 = j \text{ and } \Delta^{(\widetilde{r}+1)}(\widetilde{x}) = d + b - r_0\}. \end{aligned}$$

We observe that if x is randomly chosen with law $\mathbb{P}$, then $\widetilde{x}$ is independent of x_0 and also follows $\mathbb{P}$. We just need to take the measure to conclude. □

If we apply finitely many times (1.19), we can express $\mu^{(r)}(d)$ as a convex combination of the measures $\mu^{(0)}$ and $\mu^{(1)}$ evaluated on particular points. We recall that $\mu^{(0)} = \delta_0$ (Dirac measure on 0). We can also compute $\mu^{(1)}$.

Lemma 1.3.2. *For every $d \in \mathbb{Z}$*

$$\mu^{(1)}(d) := \begin{cases} \frac{1}{b^k} - \frac{1}{b^{k+1}} & \text{if } d = 1 - k(b-1) \text{ for some } k \in \mathbb{N} \\ 0 & \text{otherwise.} \end{cases}$$

Proof. We use again the notation $a_k^{(1)} = 1 - k(b-1)$.
By (1.12), it is trivial that if $d \neq a_k^{(1)}$ for all $k \in \mathbb{N}$ then $\mu^{(1)}(d) = 0$. Otherwise, we recall again by (1.12) that for all $k \in \mathbb{N}$, $\Delta^{(1)}(x) = a_k^{(1)}$ if and only if the right-hand side of the expansion of x is exactly a block of $(b-1)$'s of length k. Thus, if $k \geq 1$

$$\begin{aligned} \mu^{(1)}\left(a_k^{(1)}\right) &= \mathbb{P}\Big(\{x \in \mathbb{X} : x_0 = \cdots = x_{k-1} = b-1 \text{ and } x_k < b-1\}\Big) \\ &= \frac{b-1}{b^{k+1}}. \end{aligned}$$

And, if $k = 0$

$$\mu^{(1)}\left(a_k^{(1)}\right) = \mathbb{P}\Big(\{x \in \mathbb{X} : x_0 < b-1\}\Big) = \frac{b-1}{b}.$$

□

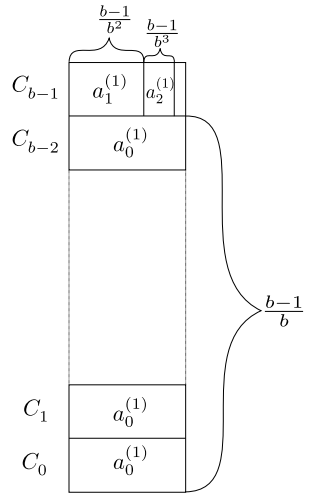


Figure 1.5: Values of $\Delta^{(1)}$ on the levels of the Rokhlin tower of order 0, and the corresponding $\mathbb{P}$-measures.

1.3.2 Inductive relation on the variance, first results

Emme and Hubert [11, Theorem 3.1] give an explicit formula for the variance in base 2. It is possible to adapt their methods in order to find a similar expression in any base. However, here we just need some basic estimations about the variance. We first deduce from Proposition 1.19 an inductive relation on the variance.

Lemma 1.3.3. *For $\widetilde{r} \in \mathbb{N}$ and $0 \leq r_0 \leq b-1$, we have the relation*

$$\mathrm{Var}(\mu^{(b\widetilde{r}+r_0)}) = \frac{b-r_0}{b}\mathrm{Var}(\mu^{(\widetilde{r})}) + \frac{r_0}{b}\mathrm{Var}(\mu^{(\widetilde{r}+1)}) + r_0(b-r_0).$$

Proof. We compute using (1.17) and (1.19)

$$\begin{aligned}\mathrm{Var}(\mu^{(b\widetilde{r}+r_0)}) &= \sum_{d\in\mathbb{Z}} d^2\mu^{(b\widetilde{r}+r_0)}(d)\\ &= \sum_{d\in\mathbb{Z}} d^2\left(\frac{b-r_0}{b}\mu^{(\widetilde{r})}(d-r_0)+\frac{r_0}{b}\mu^{(\widetilde{r}+1)}(d+b-r_0)\right)\\ &= \frac{b-r_0}{b}\sum_{d'\in\mathbb{Z}}(d'+r_0)^2\mu^{(\widetilde{r})}(d')+\frac{r_0}{b}\sum_{d'\in\mathbb{Z}}(d'+b-r_0)^2\mu^{(\widetilde{r}+1)}(d')\end{aligned}$$

Using again (1.17), we get

$$\begin{aligned}\mathrm{Var}(\mu^{(b\widetilde{r}+r_0)}) &= \frac{b-r_0}{b}\left(\mathrm{Var}(\mu^{(\widetilde{r})})+r_0^2\right)+\frac{r_0}{b}\left(\mathrm{Var}(\mu^{(\widetilde{r}+1)})+(b-r_0)^2\right)\\ &= \frac{b-r_0}{b}\mathrm{Var}(\mu^{(\widetilde{r})})+\frac{r_0}{b}\mathrm{Var}(\mu^{(\widetilde{r}+1)})+r_0(b-r_0).\end{aligned}$$

□

Even if we do not need an explicit formula of $\mathrm{Var}(\mu^{(r)})$ for a general $r\in\mathbb{N}$, we will need one in some specific cases. First, we are interested in the variance of $\mu^{(r)}$ when the expansion of r is one digit long.

Lemma 1.3.4. *If $0\le r\le b-1$ then*

$$\mathrm{Var}(\mu^{(r)}) = r(1+b-r).$$

Proof. We recall that $\mu^{(0)}=\delta_0$ so $\mathrm{Var}(\mu^{(0)})=0$. Then, if $r=1$, Lemma 1.3.3 gives

$$\mathrm{Var}(\mu^{(1)}) = \frac{b-1}{b}\mathrm{Var}(\mu^{(0)})+\frac{1}{b}\mathrm{Var}(\mu^{(1)})+(b-1).$$

It follows

$$\mathrm{Var}(\mu^{(1)}) = b.$$

Finally, if $2\le r\le b-1$, again from Lemma 1.3.3, we have

$$\mathrm{Var}(\mu^{(r)}) = \frac{b-r}{b}\mathrm{Var}(\mu^{(0)})+\frac{r}{b}\mathrm{Var}(\mu^{(1)})+r(b-r) = r(1+b-r).$$

□

Now, we are interested in the variance of $\mu^{(r)}$ when the expansion of r has a rightmost block of $(b-1)$'s of length $m\ge 1$ that is to say when there exists $\widehat{r}\in\mathbb{N}$ such that $r=b^m\widehat{r}+b^m-1$.

Lemma 1.3.5. *For $\widehat{r}\in\mathbb{N}$ and $m\ge 1$, the variance of $\mu^{(b^m\widehat{r}+b^m-1)}$ is*

$$\mathrm{Var}(\mu^{(b^m\widehat{r}+b^m-1)}) = \frac{1}{b^m}\mathrm{Var}(\mu^{(\widehat{r})})+\left(1-\frac{1}{b^m}\right)\mathrm{Var}(\mu^{(\widehat{r}+1)})+b-\frac{1}{b^{m-1}}.$$

Remark 3. *We observe that, if we take $\widehat{r} = 0$, then we find the variance of $\mu^{(r)}$ where the expansion of r is composed of only one block of $(b-1)$'s. Thus, with Lemma 1.3.4, we now have the exact value of the variance of $\mu^{(r)}$ when the expansion of r is composed of exactly 1 non-zero block:*

$$\mathrm{Var}(\mu^{(r)}) = \begin{cases} r(1+b-r) & \text{if } r = 1, \cdots, b-2, \\ 2b - \frac{2}{b^{m-1}} & \text{if } r = b^m - 1 \ (m \geq 1). \end{cases}$$

Proof. We prove the lemma by induction on $m \geq 1$.

1. If $m = 1$ then, from Lemma 1.3.3

$$\mathrm{Var}(\mu^{(b\widehat{r}+b-1)}) = \frac{1}{b}\mathrm{Var}(\mu^{(\widehat{r})}) + \frac{b-1}{b}\mathrm{Var}(\mu^{(\widehat{r}+1)}) + (b-1).$$

   That is what we want.

2. If we assume that the lemma is true for $m-1$ then we consider the integer $b^m\widehat{r} + b^m - 1$. We observe the trivial identity

$$b^m\widehat{r} + b^m - 1 = b \times \left(b^{m-1}\widehat{r} + \left(b^{m-1} - 1\right)\right) + (b-1)$$

   It follows from Lemma 1.3.3

$$\mathrm{Var}(\mu^{(b^m\widehat{r}+b^m-1)}) = \frac{1}{b}\mathrm{Var}\left(\mu^{\left(b^{m-1}\widehat{r}+\left(b^{m-1}-1\right)\right)}\right) + \frac{b-1}{b}\mathrm{Var}(\mu^{(b^{m-1}(\widehat{r}+1))}) + (b-1).$$

   We use the induction hypothesis and the fact that $\mu^{(b^{m-1}(\widehat{r}+1))} = \mu^{(\widehat{r}+1)}$ (Lemma 1.19)

$$\mathrm{Var}\left(\mu^{\left(b^{m-1}\widehat{r}+\left(b^{m-1}-1\right)\right)}\right) = \frac{1}{b^{m-1}}\mathrm{Var}(\mu^{(\widehat{r})}) + \left(1 - \frac{1}{b^{m-1}}\right)\mathrm{Var}(\mu^{(\widehat{r}+1)}) + b - \frac{1}{b^{m-2}}.$$

   Combining both gives the result for m.

□

Finally, we consider the case where the expansion of r has a units digit 1, possibly with a block of 0's on its left. This corresponds to the existence of $\widehat{r} \in \mathbb{N}$ and $m \geq 1$ such that $r = b^m\widehat{r} + 1$.

Lemma 1.3.6. *For $\widehat{r} \in \mathbb{N}$ and $m \geq 1$, the variance of $\mu^{(b^m\widehat{r}+1)}$ is*

$$\mathrm{Var}(\mu^{(b^m\widehat{r}+1)}) = \left(1 - \frac{1}{b^m}\right)\mathrm{Var}(\mu^{(\widehat{r})}) + \frac{1}{b^m}\mathrm{Var}(\mu^{\widehat{r}+1)}) + b - \frac{1}{b^{m-1}}.$$

Proof. We show by induction on $m \geq 1$.

1. If $m = 1$ then, from Lemma 1.3.3

$$\mathrm{Var}(\mu^{(b\widehat{r}+1)}) = \frac{b-1}{b}\mathrm{Var}(\mu^{(\widehat{r})}) + \frac{1}{b}\mathrm{Var}(\mu^{(\widehat{r}+1)}) + (b-1).$$

   That is exactly what we want.

2. If we assume that the formula is true for $m-1$, then we consider the integer $b^m\widehat{r}+1$. We have again by Lemma 1.3.3

$$\begin{aligned}\mathrm{Var}(\mu^{(b^m\widehat{r}+1)}) &= \mathrm{Var}(\mu^{(b\times b^{m-1}\widehat{r}+1)}) \\ &= \frac{b-1}{b}\mathrm{Var}(\mu^{(b^{m-1}\widehat{r})}) + \frac{1}{b}\mathrm{Var}(\mu^{(b^{m-1}\widehat{r}+1)}) + (b-1).\end{aligned}$$

We use the induction hypothesis.

$$\mathrm{Var}(\mu^{(b^{m-1}\widehat{r}+1)}) = \left(1-\frac{1}{b^{m-1}}\right)\mathrm{Var}(\mu^{(\widehat{r})}) + \frac{1}{b^{m-1}}\mathrm{Var}(\mu^{(\widehat{r}+1)}) + b - \frac{1}{b^{m-2}}.$$

Combining both gives the result for m.

□

1.3.3 Upper and lower bound of the variance

Since $\mathrm{Var}(\mu^{(0)}) = 0$, the case $r = 0$ is irrelevant and we suppose $r \geq 1$. We wish to find an upper and a lower bound of the variance of $\mu^{(r)}$ depending on $\rho(r)$, the number of blocks of r defined in Definition 1.1.1. For convenient reasons, it is better to think in terms of non-zero blocks. We define, for an integer r, the quantity $\lambda(r)$ which corresponds to the number of non-zero blocks.

Example 1.3.7. *We use again the example of Figure 1.1.*

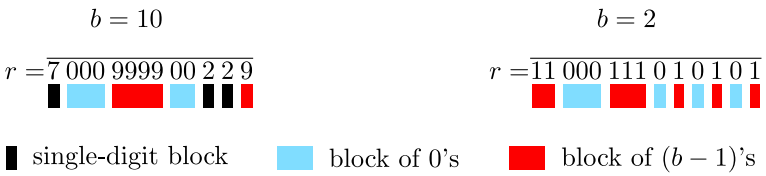


Figure 1.6: On the left-hand side, $\rho(r) = 7$ and $\lambda(r) = 5$. On the right-hand side, $\rho(r) = 9$ and $\lambda(r) = 5$.

Of course, there is a relation between $\lambda(r)$ and $\rho(r)$ (we recall that $r \geq 1$):

$$\lambda(r) \leq \rho(r) \leq 2\lambda(r). \tag{1.20}$$

We first give an upper bound of $\mathrm{Var}(\mu^{(r)})$ depending on $\lambda(r)$.

Proposition 1.3.8. *For any $r \geq 1$*

$$\mathrm{Var}(\mu^{(r)}) \leq b^2\lambda(r). \tag{1.21}$$

We need the following lemma.

Lemma 1.3.9. *For $r \geq 1$, we have the following inequality*

$$|\mathrm{Var}(\mu^{(r+1)}) - \mathrm{Var}(\mu^{(r)})| \leq b.$$

Proof. We recall that $r_0 \in \{0, \cdots, b-1\}$ is the units digit of r and the relation $r = b\widetilde{r} + r_0$. Let us prove

$$|\mathrm{Var}(\mu^{(r+1)}) - \mathrm{Var}(\mu^{(r)})| \leq \frac{|\mathrm{Var}(\mu^{(\widetilde{r}+1)}) - \mathrm{Var}(\mu^{(\widetilde{r})})|}{b} + b - 1.$$

Indeed, there are two cases.

1. If $r_0 = b-1$, then $r = b\widetilde{r} + (b-1)$ and using twice Lemma 1.3.3, we get

$$\begin{aligned}\mathrm{Var}(\mu^{(r+1)}) - \mathrm{Var}(\mu^{(r)}) &= \mathrm{Var}(\mu^{(b(\widetilde{r}+1))}) - \mathrm{Var}(\mu^{(b\widetilde{r}+b-1)}) \\ &= \frac{\mathrm{Var}(\mu^{(\widetilde{r}+1)}) - \mathrm{Var}(\mu^{(\widetilde{r})})}{b} - (b-1).\end{aligned}$$

2. If $r_0 \in \{0, \cdots, b-2\}$ then, using the same tools, we compute

$$\begin{aligned}\mathrm{Var}(\mu^{(r+1)}) - \mathrm{Var}(\mu^{(r)}) &= \mathrm{Var}(\mu^{(b\widetilde{r}+r_0+1)}) - \mathrm{Var}(\mu^{(b\widetilde{r}+r_0)}) \\ &= \frac{\mathrm{Var}(\mu^{(\widetilde{r}+1)}) - \mathrm{Var}(\mu^{(\widetilde{r})})}{b} + b - 2r_0 - 1.\end{aligned}$$

   We observe that $|b - 2r_0 - 1| \leq b - 1$.

Then, we conclude by an easy induction on the number of digits of r and by checking that

$$\begin{aligned}|\mathrm{Var}(\mu^{(r+1)}) - \mathrm{Var}(\mu^{(r)})| &\leq \frac{|\mathrm{Var}(\mu^{(\widetilde{r}+1)}) - \mathrm{Var}(\mu^{(\widetilde{r})})|}{b} + b - 1 \\ &\leq \frac{b}{b} + b - 1 = b.\end{aligned}$$

□

Proof of Proposition 1.3.8. Observe that we have

$$b^2 \geq j(1+b-j) \quad \text{for all } j = 1, \cdots b-1, \tag{C_0}$$

$$b^2 \geq 2b, \tag{C_1}$$

$$b^2 \geq jb \quad \text{for all } j = 0, \cdots, b-1. \tag{C_2}$$

We proceed by induction on $\lambda(r) \geq 1$.

1. Initialisation: if $\lambda(r) = 1$ then we have two cases depending on the type of blocks we are considering. However, from Lemma 1.3.4, Lemma 1.3.5 and Conditions (C_0) and (C_1), we can deduce that in both cases we have $\mathrm{Var}(\mu^{(r)}) \leq b^2$.

2. Inductive step: we let $n > 1$ and we assume that if $\lambda(r) \leq n$ then $\mathrm{Var}(\mu^{(r)}) \leq b^2\lambda(r)$. We now assume that our $r \in \mathbb{N}$ satisfies $\lambda(r) = n+1$. Let $\ell \geq 0$, $\widehat{r} \in \mathbb{N}$ and B_1 a non-zero block such that we can write the expansion of r as follow

$$\overline{\widehat{r}\ B_1\ 0^\ell}.$$

   We can assume $\ell = 0$ but $\widehat{r}$ may have a rightmost block composed of 0's. We discuss on the type of B_1.

(a) If B_1 is a block of $(b-1)$'s of length m then $r = b^m\widehat{r} + b^m - 1$ and we have the trivial equality

$$\begin{aligned}\mathrm{Var}(\mu^{(r)}) = \mathrm{Var}(\mu^{(\widehat{r})}) &+ \Big(\mathrm{Var}(\mu^{(r)}) - \mathrm{Var}(\mu^{(r+1)})\Big) \\ &- \Big(\mathrm{Var}(\mu^{(\widehat{r})}) - \mathrm{Var}(\mu^{(\widehat{r}+1)})\Big).\end{aligned}$$

Using Lemma 1.3.9, the inductive hypothesis and Condition (C_1)

$$\mathrm{Var}(\mu^{(r)}) \leq b^2 n + b + b \leq b^2(n+1).$$

(b) If B_1 is a single-digit block r_0 then $r = b\widehat{r} + r_0$ and with we have another trivial equality

$$\mathrm{Var}(\mu^{(r)}) = \mathrm{Var}(\mu^{(\widehat{r})}) + \sum_{k=1}^{r_0} \mathrm{Var}(\mu^{(b\widehat{r}+k)}) - \mathrm{Var}(\mu^{(b\widehat{r}+k-1)}).$$

Then, using Lemma 1.3.9, the inductive hypothesis and Condition (C_2), we find

$$\mathrm{Var}(\mu^{(r)}) \leq b^2 n + r_0 b \leq b^2(n+1).$$

This concludes the inductive step and the proof.

□

We also have a lower bound depending on $\lambda(r)$.

Proposition 1.3.10. *For any $r \geq 1$*

$$\mathrm{Var}(\mu^{(r)}) \geq \frac{b}{4}\lambda(r). \tag{1.23}$$

Proof. Observe that

$$\frac{b}{4} \leq \min\{j(b-j) : j = 1, \cdots b-1\}, \tag{C_3}$$

$$\frac{b}{4} \leq (b-1), \tag{C_4}$$

$$\frac{b}{4} \leq \min\{j(b-j-\frac{1}{2}) : j = 1, \cdots b-1\}, \tag{C_5}$$

Of course, some of these conditions are redundant but we keep them all for simplicity because each one will be used in the proof. We prove the result using induction on $\lambda(r) \geq 1$.

1. Initialisation: if $\lambda(r) = 1$ then we have two cases depending on the type of blocks we are considering. However, from Lemma 1.3.4, Lemma 1.3.5 and Conditions (C_3) and (C_4), we can deduce $\mathrm{Var}(\mu^{(r)}) \geq \frac{b}{4}$.

2. Inductive step: we let $n > 1$ and we assume that if $\lambda(r) \leq n$ then $\mathrm{Var}(\mu^{(r)}) \geq \frac{b}{4}\lambda(r)$. We now take $r \in \mathbb{N}$ such that $\lambda(r) = n+1$, and we write

$$r = b\widetilde{r} + r_0.$$

Without loss of generality, we can assume that $r_0 \neq 0$. Indeed, if $r_0 = 0$ then $\mathrm{Var}(\mu^{(r)}) = \mathrm{Var}(\mu^{(\widetilde{r})})$ and $\lambda(r) = \lambda(\widetilde{r})$.

From Lemma 1.3.3, we get

$$\mathrm{Var}(\mu^{(r)}) = \frac{b-r_0}{b}\mathrm{Var}(\mu^{(\widetilde{r})}) + \frac{r_0}{b}\mathrm{Var}(\mu^{(\widetilde{r}+1)}) + r_0(b-r_0). \tag{1.25}$$

We discuss about the value of $\lambda(\widetilde{r})$ which can be either n or $n+1$.

(a) If $\lambda(\widetilde{r}) = n$, which means that $r_0 \neq b-1$ or $r_1 \neq b-1$: we use the induction hypothesis to get $\mathrm{Var}(\mu^{(\widetilde{r})}) \geq \frac{b}{4}\lambda(\widetilde{r}) = \frac{bn}{4}$, and it follows that

$$\mathrm{Var}(\mu^{(r)}) \geq \frac{b-r_0}{4}n + \frac{r_0}{b}\mathrm{Var}(\mu^{(\widetilde{r}+1)}) + r_0(b-r_0).$$

We now observe that $n-2 \leq \lambda(\widetilde{r}+1) \leq n+1$. The reader is referred to Figure 1.7 (with $\widetilde{r}$ instead of r) for more details. We consider two cases.

i. If $\lambda(\widetilde{r}+1) = n+1$ then we cannot apply the induction hypothesis. In this case $\widetilde{r}$ is a multiple of b if $b \geq 3$ and even b^2 when $b = 2$ (see Figure 1.7). So, there exists $\widehat{r} \in \mathbb{N}$ and $m \geq 1$ (or 2 if $b = 2$) such that

$$\widetilde{r} = b^m\widehat{r} \qquad \text{and} \qquad b \not| \widehat{r}.$$

We can apply Lemma 1.3.6

$$\mathrm{Var}(\mu^{(\widetilde{r}+1)}) = \left(1 - \frac{1}{b^m}\right)\mathrm{Var}(\mu^{(\widehat{r})}) + \frac{1}{b^m}\mathrm{Var}(\mu^{(\widehat{r}+1)}) + b - \frac{1}{b^{m-1}}.$$

Since $\mu^{(\widetilde{r})} = \mu^{(\widehat{r})}$ and $\lambda(\widetilde{r}) = n$, we deduce from the induction hypothesis and (1.25) that

$$\begin{aligned}\mathrm{Var}(\mu^{(r)}) &\geq \frac{b-r_0}{4}n + \frac{r_0}{b}\left[\frac{b}{4}n + b - \frac{3}{2b^{m-1}}\right] + \frac{b}{4} \\ &\geq \frac{b}{4}(n+1) + r_0\left(1 - \frac{3}{2b^{m-1}}\right) \\ &\geq \frac{b}{4}(n+1).\end{aligned}$$

ii. If $\lambda(\widetilde{r}+1) \neq n+1$ then we can apply the induction hypothesis.

$$\begin{aligned}\mathrm{Var}(\mu^{(r)}) &\geq \frac{b-r_0}{4}n + \frac{r_0}{2}(n-2) + r_0(b-r_0) \\ &\geq \frac{b}{4}n - \frac{r_0}{2} + r_0(b-r_0).\end{aligned}$$

Thanks to (C_5)

$$r_0(b-r_0) - \frac{r_0}{2} \geq \frac{b}{4}.$$

And so

$$\mathrm{Var}(\mu^{(r)}) \geq \frac{b}{4}n + \frac{b}{4} = \frac{b}{4}(n+1).$$

We conclude the case $\lambda(\widetilde{r}) = n$.

(b) If $\lambda(\widetilde{r}) = n+1$: it means that the rightmost block in the expansion of r is a block of $(b-1)$'s of length $m \geq 2$. So, there exists $\widehat{r} \in \mathbb{N}$ such that $r = b^m\widehat{r} + b^m - 1$ and the rightmost digit in the expansion of $\widehat{r}$ is not $(b-1)$, that is to say $b \not| \widehat{r}+1$. We are in the context of Lemma 1.3.5, we have

$$\mathrm{Var}(\mu^{(r)}) = \frac{1}{b^m}\mathrm{Var}(\mu^{(\widehat{r})}) + \left(1 - \frac{1}{b^m}\right)\mathrm{Var}(\mu^{(\widehat{r}+1)}) + b - \frac{1}{b^{m-1}}.$$

Since $\lambda(\widehat{r}) = n$ and $n-1 \leq \lambda(\widehat{r}+1) \leq n+1$ ($\widehat{r}$ does not start with a block of $(b-1)$'s so $\lambda(\widehat{r}+1) \neq n-2$, see Figure 1.7), we have now

$$\mathrm{Var}(\mu^{(r)}) \geq \frac{n}{4b^{m-1}} + \left(1 - \frac{1}{b^m}\right)\mathrm{Var}(\mu^{(\widehat{r}+1)}) + b - \frac{1}{b^{m-1}}. \tag{1.26}$$

We discuss about the possible values of $\lambda(\widehat{r}+1)$.

i. If $\lambda(\widehat{r}+1) = n+1$ then it is just as in the point (a)i. of this proof, it means that $\widehat{r}$ starts with a block of 0's. So we let $m' \geq 1$ (2 if $b = 2$) and $\widehat{\widehat{r}} \in \mathbb{N}$ such that $\widehat{r} = b^{m'}\widehat{\widehat{r}} =$ and $b \not| \widehat{\widehat{r}}$. We are again in the context of Lemma 1.3.6

$$\mathrm{Var}(\mu^{(\widehat{r}+1)}) = \left(1 - \frac{1}{b^{m'}}\right)\mathrm{Var}(\mu^{(\widehat{\widehat{r}})}) + \frac{1}{b^{m'}}\mathrm{Var}(\mu^{(\widehat{\widehat{r}}+1)}) + b - \frac{1}{b^{m'-1}}.$$

We observe $\lambda(\widehat{\widehat{r}}) = \lambda(\widehat{r}) = n$ and $n-2 \leq \lambda(\widehat{\widehat{r}}+1) \leq n$ (again, it cannot be $n+1$ because $\widehat{\widehat{r}}$ does not start with a block of 0's).
So we have

$$\begin{aligned}\mathrm{Var}(\mu^{(\widehat{r}+1)}) &\geq \left(1 - \frac{1}{b^{m'}}\right)\frac{b}{4}n + \frac{1}{4b^{m'-1}}(n-2) + b - \frac{1}{b^{m'-1}} \\ &= \frac{b}{4}n + b - \frac{3}{2b^{m'-1}}.\end{aligned}$$

From (1.26) we deduce

$$\mathrm{Var}(\mu^{(r)}) \geq \frac{b}{4}n + \left(1 - \frac{1}{b^m}\right)\left(b - \frac{3}{2b^{m'-1}}\right) + b - \frac{1}{b^{m-1}}.$$

We observe that

$$b - \frac{3}{2b^{m'-1}} \geq 0$$

as well as

$$b - \frac{1}{b^{m-1}} \geq \frac{b}{4}.$$

So we can write

$$\mathrm{Var}(\mu^{(r)}) \geq \frac{b}{4}(n+1).$$

ii. If $\lambda(\widehat{r}+1) \neq n+1$ then we can apply the induction hypothesis.

$$\begin{aligned}\mathrm{Var}(\mu^{(r)}) &\geq \frac{n}{4b^{m-1}} + \left(1 - \frac{1}{b^m}\right)\frac{b}{4}(n-1) + b - \frac{1}{b^{m-1}} \\ &\geq \frac{b}{4}n - \left(1 - \frac{1}{b^m}\right)\frac{b}{4} + b - \frac{1}{b^{m-1}}\end{aligned}$$

We observe that

$$b - \frac{1}{b^{m-1}} - \left(1 - \frac{1}{b^m}\right)\frac{b}{4} \geq \frac{b}{4},$$

so

$$\mathrm{Var}(\mu^{(r)}) \geq \frac{b}{4}n + \frac{b}{4} = \frac{b}{4}(n+1).$$

It concludes the case $\lambda(\widetilde{r}) = n+1$. The statement is thus true when $\lambda(r) = n+1$.

□

The following figure shows how the number of blocks behaves of an integer when we add 1 to it.

$b = 2$

Rightmost block(s) of r	$\lambda(r+1) - \lambda(r)$
	1 if length( )≥ 2 0 if length( )$= 1$
	0 if length( )≥ 2 -1 if length( )$= 1$

$b \geq 3$

Rightmost block(s) of r	$\lambda(r+1) - \lambda(r)$
	1
	0
	-1 if $\blacksquare = b-2$ 0 otherwise
	0
	0
	-1
	-1
	-2 if $\blacksquare = b-2$ -1 otherwise


block of 0's


single digit between 1 and $b-2$


block of $(b-1)$'s

Figure 1.7: Variations of the number of non-zero blocks when we add 1.

Proof of Theorem 1.1.2. We use Proposition 1.3.8, Proposition 1.3.10 and (1.20) to get the result. □

1.4 A ϕ-mixing process

We work on the probability space $(\mathbb{X}, \mathscr{B}(\mathbb{X}), \mathbb{P})$. For a given integer r, $\Delta^{(r)}$ is viewed as a random variable with law $\mu^{(r)}$ by Corollary 1.2.2 (the randomness comes from the argument x of $\Delta^{(r)}$, considered as a random outcome in $\mathbb{X}$ with law $\mathbb{P}$). Our purpose in this section is to study the asymptotic behaviour of $\mu^{(r)}$ as the number of blocks $\rho(r)$ goes to infinity. For this, we will decompose $\Delta^{(r)}$ as a sum

$$\Delta^{(r)} = \sum_{i=1}^{\lambda(r)} X_i^{(r)}$$

where $\lambda(r)$ is the number of non-zero blocks in the base-b expansion of r (see Section 1.3.3), and $(X_1^{(r)}, \cdots, X_{\lambda(r)}^{(r)})$ is a finite process defined in the next section. We will prove and use some mixing properties of this process to get our result.

1.4.1 The process

Again, the case $r = 0$ is irrelevant so we assume $r \geq 1$. For $1 \leq i \leq \lambda(r)$, we will write B_i as the i^{th} non-zero block present in the expansion of r, starting from the left-hand side of the expansion and ending at the units digit. We now define $r[i]$ as the integer whose base-b expansion is obtained as follows: for $k = i+1, \cdots, \lambda(r)$, we replace the block B_k by a block of 0's of the same length (see Figure 1.8 below). We observe that $r[\lambda(r)] = r$.

$$\begin{aligned}
r &= \overline{\underbrace{7}_{B_1}\ 0\,0\,0\ \underbrace{9\,9\,9}\ \underbrace{9}_{B_2}\ 0\,0\ \underbrace{2}_{B_3}\ \underbrace{2}_{B_4}\ \underbrace{9}_{B_5}} \\
r[1] &= \overline{7\ \ 0\,0\,0\ \ 0\,0\,0\ \ 0\,0\,0\ \ 0\,0\,0} \\
r[2] &= \overline{7\ \ 0\,0\,0\ \ 9\,9\,9\ \ 9\,0\,0\ \ 0\,0\,0} \\
r[3] &= \overline{7\ \ 0\,0\,0\ \ 9\,9\,9\ \ 9\,0\,0\ \ 2\,0\,0} \\
r[4] &= \overline{7\ \ 0\,0\,0\ \ 9\,9\,9\ \ 9\,0\,0\ \ 2\,2\,0} \\
r[5] &= r
\end{aligned}$$

Figure 1.8: Example in base $b = 10$.

With the convention $r[0] := 0$, we observe the trivial equality

$$r = \sum_{i=1}^{\lambda(r)} r[i] - r[i-1].$$

For $1 \leq i \leq \lambda(r)$, we define almost everywhere on $\mathbb{X}$ (see Subsection 1.2.2)

$$X_i^{(r)} := \Delta^{(r[i]-r[i-1])} \circ T^{r[i-1]}.$$

Since $r[i]-r[i-1] = \overline{B_i 0 \cdots 0}$, the function $X_i^{(r)}$ is a random variable corresponding to the action of the i^{th} block B_i once the previous blocks have already been taken into consideration. From (1.10), we deduce

$$\Delta^{(r)} = \sum_{i=1}^{\lambda(r)} X_i^{(r)}.$$

In particular, if $x \in \mathbb{X}$ is randomly chosen with law $\mathbb{P}$, then $\displaystyle\sum_{i=1}^{\lambda(r)} X_i^{(r)}(x)$ follows the law $\mu^{(r)}$. Hence, the standard deviation σ_r of $\mu^{(r)}$ defined in Theorem 1.1.4 satisfies

$$\sigma_r^2 = \mathrm{Var}\left(\sum_{i=1}^{\lambda(r)} X_i^{(r)}\right).$$

We first show that every moment of $X_i^{(r)}$ is bounded from above by a constant independent of r and i.

Lemma 1.4.1. *For every $k \in \mathbb{N}$, there exists a constant $C_k > 0$ such that*

$$\forall r \in \mathbb{N}, \forall 1 \leq i \leq \lambda(r), \qquad \mathbb{E}\left(\left|X_i^{(r)}\right|^k\right) \leq C_k.$$

Proof. Let $k \in \mathbb{N}$. If $X_i^{(r)}$ corresponds to the action of a single-digit block, that is the action of one digit α between 1 and $b-2$, then its law is given by $\mu^{(\alpha)}$ whose moments are all finite (see Corollary 1.2.2). So, we have in this case

$$\mathbb{E}\left(\left|X_i^{(r)}\right|^k\right) \leq \max\left\{\mathbb{E}\left(\left|\Delta^{(\alpha)}\right|^k\right) : 1 \leq \alpha \leq b-2\right\}.$$

Now, if $X_i^{(r)}$ corresponds to the action of a block of $(b-1)$'s, there exist two integers $n \leq m$ such that $r[i] - r[i-1] = b^m - b^n$. It follows from the definition of $X_i^{(r)}$ and (1.10) that

$$X_i^{(r)} = \Delta^{(b^m-b^n)} \circ T^{b^n} \overset{d}{=} \Delta^{(b^m-b^n)} = \Delta^{(b^m)} - \Delta^{(b^n)} \circ T^{b^m-b^n}$$

where $\overset{d}{=}$ means the equality in distribution. Since $\mu^{(b^n)} = \mu^{(b^m)} = \mu^{(1)}$, we can write $X_i^{(r)}$ as the difference of two dependent random variables following the law $\mu^{(1)}$ which has finite moments. So there exists a constant that depends only on k such that $\mathbb{E}\left(|X_i^{(r)}|^k\right)$ is bounded by this constant. □

The next part is devoted to the estimation of the so-called ϕ-mixing coefficients for the finite sequence $(X_i^{(r)})_{1 \leq i \leq \lambda(r)}$.

1.4.2 The ϕ-mixing coefficients

There exist many types of mixing coefficients (see e.g. the survey [5] by Bradley). Those we are working with are commonly called "ϕ-mixing coefficients".

Definition 1.4.2. *Let $(X_i)_{i\geq 1}$ be a (finite or infinite) sequence of random variables. The associated* ϕ-mixing coefficients $\phi(k)$, $k \geq 1$, *are defined by*

$$\phi(k) := \sup_{p\geq 1} \sup_{A,B} |\mathbb{P}_A(B) - \mathbb{P}(B)|$$

where the second supremum is taken over all events A and B such that

- $A \in \sigma(X_i : 1 \leq i \leq p)$,
- $\mathbb{P}(A) > 0$ *and*
- $B \in \sigma(X_i : i \geq k + p)$.

By convention, if X_i is not defined when $i \geq k + p$ then the σ-algebra is trivial.

In the case of a finite sequence $(X_1, \cdots, X_n)$, the convention implies that $\phi(k) = 0$ for $k \geq n$. We now give an upper bound on the ϕ-mixing coefficients for the process $(X_i^{(r)})$ defined in Section 1.4.1.

Lemma 1.4.3. *For $r \geq 1$, the mixing coefficients of $(X_i^{(r)})_{1\leq i\leq \lambda(r)}$ satisfy*

$$\forall k \geq 1, \quad \phi(k) \leq 2\left(\frac{b-1}{b}\right)^{\frac{k}{2}-1}.$$

Proof. Let k and p be two integers. We observe that if $k = 1, 2$, the inequality is trivial so we assume that $k \geq 3$. We call *buffer strip* the set of indices corresponding to the positions of the digits between B_p and B_{k+p} (both excluded). It depends on r, p and k so we denote it by $\mathcal{I}_{r,p,k}$. We consider the event

$$C := \{x \in \mathbb{X} : \exists j \in \mathcal{I}_{r,p,k} \text{ with } r_j < b - 1 \text{ such that } x_j = 0\}.$$

We are going to show that, for k large enough, C is a high-probability event and that, conditionned to C, two events $A \in \sigma\left(X_i^{(r)} : 1 \leq i \leq p\right)$ and $B \in \sigma\left(X_i^{(r)} : i \geq k + p\right)$ are always independent.

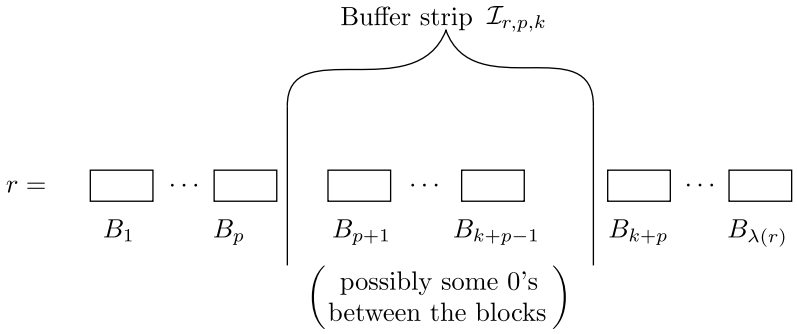


Figure 1.9: Visualization of the buffer strip.

Denoting by $\overline{C}$ the complement of C in $\mathbb{X}$, we have

$$\mathbb{P}(\overline{C}) = \left(\frac{b-1}{b}\right)^t \tag{1.27}$$

where $t := \Big|\{j \in \mathcal{I}_{r,p,k} \,:\, r_j \neq b-1\}\Big|$. We are going to show that

$$t \geq \frac{k}{2} - 1. \tag{1.28}$$

Indeed, there are $k-1$ non-zero blocks in the buffer strip. Let $\ell \in \mathbb{N}$ be the number of blocks of $(b-1)$'s. Then there are $k-1-\ell$ single-digit blocks. There are two cases.

1. If $\ell \leq \frac{k}{2}$, then $k-1-\ell \geq \frac{k}{2}-1$ and we get (1.28).
2. Otherwise, $\ell > \frac{k}{2}$. Since the blocks of $(b-1)$'s are separated using blocks of zeros or single-digit blocks, there are at least $\frac{k}{2}-1$ blocks of zeros of single-digit blocks, which also yields (1.28).

So, we get from (1.27) and (1.28) that

$$\mathbb{P}(C) \geq 1 - \left(\frac{b-1}{b}\right)^{\frac{k}{2}-1} > 0.$$

Now, let $A \in \sigma\left(X_i^{(r)} : 1 \leq i \leq p\right)$ with $\mathbb{P}(A) > 0$ and $B \in \sigma\left(X_i^{(r)} : i \geq k+p\right)$.

Observe that A and C are independent. Indeed, C only depends on indices in $\mathcal{I}_{r,p,k}$ while A, by construction of the random variables $(X_i^{(r)})_{1\leq i\leq \lambda(r)}$, only depends on the subset of indices on the left-hand side of the buffer strip. We deduce

$$\mathbb{P}(A \cap C) = \mathbb{P}(A)\mathbb{P}(C) > 0. \tag{1.29}$$

Observe also that, conditionned to C, A and B are independent. Indeed, when C is realized, there exists an index j in $\mathcal{I}_{r,p,k}$ such that $r_j \neq b-1$ and $x_j = 0$. At this position, a carry cannot be created and, furthermore, a carry propagation coming from the right-hand side will be stopped at this index. In other words, when C is realized, the carries created by the blocks B_i, $i \geq k+p$, never spread on the left-hand side of the buffer strip. Moreover, we deduce that A and $B \cap C$ are independent. Indeed, we compute

$$\begin{aligned}\mathbb{P}(A \cap B \cap C) &= \mathbb{P}_C(A \cap B)\mathbb{P}(C) \\ &= \mathbb{P}_C(A)\mathbb{P}_C(B)\mathbb{P}(C) \\ &= \mathbb{P}(A)\mathbb{P}(B \cap C).\end{aligned} \tag{1.30}$$

Now, we have

$$|\mathbb{P}_A(B) - \mathbb{P}(B)| \leq |\mathbb{P}_A(B) - \mathbb{P}_{A\cap C}(B)| + |\mathbb{P}_{A\cap C}(B) - \mathbb{P}(B)|. \tag{1.31}$$

But, using (1.29) and (1.30), we obtain

$$\mathbb{P}_{A\cap C}(B) = \frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(A \cap C)} = \frac{\mathbb{P}(B \cap C)}{\mathbb{P}(C)} = \mathbb{P}_C(B).$$

So, (1.31) becomes

$$|\mathbb{P}_A(B) - \mathbb{P}(B)| \leq |\mathbb{P}_A(B) - \mathbb{P}_{A\cap C}(B)| + |\mathbb{P}_C(B) - \mathbb{P}(B)|. \tag{1.32}$$

Now, observe that for any event D, by the law of total probability,

$$\begin{aligned}\left|\mathbb{P}(D) - \mathbb{P}_C(D)\right| &= \left|\mathbb{P}(C)\mathbb{P}_C(D) + \mathbb{P}\left(\overline{C}\right)\mathbb{P}_{\overline{C}}(D) - \mathbb{P}_C(D)\right| \\ &= \mathbb{P}\left(\overline{C}\right)\left|\mathbb{P}_{\overline{C}}(D) - \mathbb{P}_C(D)\right|.\end{aligned}$$

Since both terms are non negative and less than 1, we have

$$\left|\mathbb{P}_{\overline{C}}(D) - \mathbb{P}_C(D)\right| \leq 1.$$

It follows

$$\left|\mathbb{P}(D) - \mathbb{P}_C(D)\right| \leq \mathbb{P}\left(\overline{C}\right).$$

These computations are also true replacing $\mathbb{P}$ by $\mathbb{P}_A$. We observe that the measure $\mathbb{P}_A$ conditionned to C is the measure $\mathbb{P}_{A\cap C}$. So, coming back to (1.32), we get

$$|\mathbb{P}_A(B) - \mathbb{P}(B)| \leq \mathbb{P}_A\left(\overline{C}\right) + \mathbb{P}\left(\overline{C}\right) = 2\mathbb{P}\left(\overline{C}\right) \leq 2\left(\frac{b-1}{b}\right)^{\frac{k}{2}-1}.$$

□

1.5 Proof of Theorem 1.1.4

1.5.1 A result from Sunklodas and first step of the proof

In this section, we state a result by Sunklodas [29] about the speed of convergence in the Central Limit Theorem for ϕ-mixing sequences. Actually, our formulation is new but the proof is an immediate consequence of [29, Theorem 1].

We first need to introduce some notations. Let Y be a standard normal random variable. Consider $\xi_1, \cdots, \xi_n$ a finite sequence of n random variables with ϕ-mixing coefficients $\phi(k)$ for $k = 1, 2, \cdots$. Write

$$V := \sqrt{\mathrm{Var}\left(\sum_{i=1}^{n}\xi_i\right)} \quad \text{and} \quad Z := \sum_{i=1}^{n}\frac{\xi_i}{V}.$$

We need to add, for technical reason, some other notations.

$$\Phi_{1/2} := \sum_{k\geq 1} k\sqrt{\phi(k)} \quad \text{and} \quad \overline{\Phi_{1/2}} := \max\left\{\sqrt{\Phi_{1/2}}, \Phi_{1/2}^2\right\}.$$

We observe that $\Phi_{1/2}$ is defined by a sum on, actually, a finite number of non-zero terms because we are in the case of a finite sequence of random variables. Our formulation of [29, Theorem 1] is the following.

Theorem 1.5.1. *Assume for $i = 1, \cdots, n$ that $\mathbb{E}(\xi_i) = 0$ and $\mathbb{E}(\xi_i^4) < \infty$. Let $h : \mathbb{R} \to \mathbb{R}$ be a thrice differentiable function such that $||h^{'''}||_\infty < \infty$. Then*

$$\left|\mathbb{E}\left(h(Z) - h(Y)\right)\right| \leq ||h^{'''}||_\infty \left(\frac{5}{2} + 28\overline{\Phi_{1/2}}\right) \sum_{i=1}^{n} \mathbb{E}\left(\left|\frac{\xi_i}{V}\right|^3\right) + 120||h^{'''}||_\infty \overline{\Phi_{1/2}} \sqrt{\sum_{i=1}^{n} \mathbb{E}\left(\left|\frac{\xi_i}{V}\right|^2\right)} \sqrt{\sum_{i=1}^{n} \mathbb{E}\left(\left|\frac{\xi_i}{V}\right|^4\right)}.$$

We are going to show that we can apply this result to prove Theorem 1.1.4. We start with (1.3).

Proof of (1.3). Let $r \in \mathbb{N}^*$. The variables $(X_i^{(r)})_{1 \leq i \leq \lambda(r)}$ are of zero-mean and have a finite moment of order 4. Lemma 1.4.3 gives a universal upper bound for $\Phi_{1/2}$. So we can apply Theorem 1.5.1 to the sequence $(X_i^{(r)})_{1 \leq i \leq \lambda(r)}$. We observe that, from (1.20), Theorem 1.1.2 and Lemma 1.4.1, for $j = 2, 3, 4$

$$\sum_{i=1}^{\lambda(r)} \mathbb{E}\left(\left|\frac{X_i^{(r)}}{\sigma_r}\right|^j\right) = \frac{C_j \lambda(r)}{\sigma_r^j} \leq \frac{2^j C_j}{b^{\frac{j}{2}}} \rho(r)^{1-\frac{j}{2}}.$$

So, by Theorem 1.5.1, there exists $K > 0$ such that

$$\left|\mathbb{E}\left(h\left(\sum_{i=1}^{\lambda(r)} \frac{X_i^{(r)}}{\sigma_r}\right) - h(Y)\right)\right| \leq \frac{K||h'''||_\infty}{\sqrt{\rho(r)}} \tag{1.33}$$

□

It remains to prove (1.4).

1.5.2 Speed of convergence of the cumulative distribution functions

Observe that for all $t \in \mathbb{R}$

$$F_r(t) - F(t) = \mathbb{E}\left(\mathbb{1}_{]-\infty,t]}\left(\sum_{i=1}^{\lambda(r)} \frac{X_i^{(r)}}{\sigma_r}\right) - \mathbb{1}_{]-\infty,t]}(Y)\right).$$

The idea is thus to find, for $t \in \mathbb{R}$, a family of thrice differentiable function $(h_{t,\varepsilon})_{\varepsilon > 0}$ with, for every $\varepsilon > 0$, $||h_{t,\varepsilon}^{'''}||_\infty < \infty$ and which converges pointwise to the indicator funtion $\mathbb{1}_{]-\infty,t]}$ when ε tends to 0.

Approximation of the indicator function

There exists a function $f : \mathbb{R} \to \mathbb{R} \in \mathcal{C}^3(\mathbb{R})$ satisfying the following conditions $f^{'}(0) = f^{'}(1) = f^{''}(0) = f^{''}(1) = f^{'''}(0) = f^{'''}(1) = 0$, $f(t) = 1$ if $t \leq 0$, $f(t) = 0$ if $t \geq 1$ and $0 \leq f(t) \leq 1$, for all

real t. Then we define the linear function $\theta_{t,\varepsilon} : [t-\varepsilon, t+\varepsilon] \to [0,1]$, $u \mapsto \frac{1}{2\varepsilon}(\varepsilon - t + u)$. Finally, we get our approximation by

$$h_{t,\varepsilon}(u) := \begin{cases} 1 & \text{if } u \leq t-\varepsilon, \\ f \circ \theta_{t,\varepsilon}(u) & \text{if } t-\varepsilon \leq u \leq t+\varepsilon, \\ 0 & \text{otherwise.} \end{cases}$$

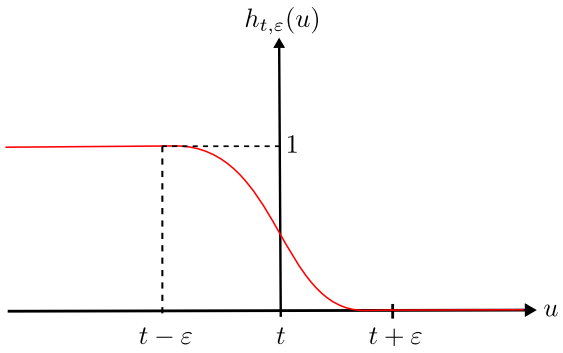


Figure 1.10: Graph of $h_{t,\varepsilon}$.

We have the following properties satisfy by $h_{t,\varepsilon}$.

Lemma 1.5.2. *Let $r \geq 1$, let Y be a standard normal random variable.*

1. *$\forall t \in \mathbb{R}$, the sequence of $\mathcal{C}^3(\mathbb{R})$ functions $(h_{t,\varepsilon})_\varepsilon$ converges pointwise to the indicator function $\mathbb{1}_{]-\infty;t]}$ when ε tends to 0.*
2. *$\forall \varepsilon > 0$, $\forall t \in \mathbb{R}$, $||h'''_{t,\varepsilon}||_\infty = \frac{||f'''||_\infty}{8\varepsilon^3}$ and, in particular, the upper bound is independent of t.*
3. *$\forall \varepsilon > 0$, we have, for any random variable X*

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}(X \leq t) - \mathbb{P}(Y \leq t) \right| \leq \sup_{t\in\mathbb{R}} \left| \mathbb{E}\left(h_{t,\varepsilon}(X) - h_{t,\varepsilon}(Y)\right) \right| + \frac{4\varepsilon}{\sqrt{2\pi}}. \tag{1.34}$$

Last step of the proof of (1.4)

Proof of (1.4). Let $\varepsilon > 0$. From (1.33) and (1.34), we obtain

$$\begin{aligned} \sup_{t\in\mathbb{R}} \left| \mathbb{P}\left(\frac{\Delta^{(r)}}{\sigma_r} \leq t\right) - \mathbb{P}(Y \leq t) \right| &\leq \sup_{t\in\mathbb{R}} \left| \mathbb{E}\left(h_{t,\varepsilon}\left(\frac{\Delta^{(r)}}{\sigma_r}\right) - h_{t,\varepsilon}(Y)\right) \right| + \frac{4\varepsilon}{\sqrt{2\pi}} \\ &\leq \sup_{t\in\mathbb{R}} \left(\frac{K||h'''_{t,\varepsilon}||_\infty}{\sqrt{\rho(r)}} \right) + \frac{4\varepsilon}{\sqrt{2\pi}} \end{aligned}$$

Lemma 1.5.2 gives $||h_{t,\varepsilon}^{'''}||_\infty = \dfrac{||f^{'''}||_\infty}{8\varepsilon^3}$. So, we obtain

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}\left(\frac{\Delta^{(r)}}{\sigma_r}\leq t\right)-\mathbb{P}\left(Y\leq t\right)\right| \leq \frac{||f^{'''}||_\infty K}{8\varepsilon^3\sqrt{\rho(r)}}+\frac{4\varepsilon}{\sqrt{2\pi}}$$

Now, we choose $\varepsilon>0$ such that $\dfrac{1}{\varepsilon^3\sqrt{\rho(r)}}=\dfrac{4\varepsilon}{\sqrt{2\pi}}$ and get the existence of a constant $\widetilde{K}>0$ such that

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}\left(\frac{\Delta^{(r)}}{\sigma_r}\leq t\right)-\mathbb{P}\left(Y\leq t\right)\right| \leq \frac{\widetilde{K}}{\rho(r)^{\frac{1}{8}}}.$$

□

It only remains to prove Lemma 1.5.2.

Proof of Lemma 1.5.2

Proof of Lemma 1.5.2. Let $\varepsilon>0$. The first point is trivial by construction of $h_{t,\varepsilon}$. The second point is also quite simple to show. We write

$$\begin{aligned}\sup_{u\in\mathbb{R}}\left|h_{t,\varepsilon}^{'''}(u)\right| &= \sup_{t-\varepsilon\leq u\leq t+\varepsilon}\left|h_{t,\varepsilon}^{'''}(u)\right| \\ &= \sup_{t-\varepsilon\leq u\leq t+\varepsilon}\left|\theta_{t,\varepsilon}^{'}(u)\right|^3\left|f^{'''}\circ\theta_{t,\varepsilon}(u)\right| \\ &= \frac{1}{8\varepsilon^3}\sup_{0\leq u\leq 1}\left|f^{'''}(u)\right|.\end{aligned}$$

For the last point, let t and $x\in\mathbb{R}$. Since

$$h_{t-\varepsilon,\varepsilon}(x)\leq \mathbb{1}_{]-\infty,t]}(x)\leq h_{t+\varepsilon,\varepsilon}(x),$$

we deduce that for any random variable X

$$\mathbb{E}\left(h_{t-\varepsilon,\varepsilon}\left(X\right)\right)\leq\mathbb{P}\left(X\leq t\right)\leq\mathbb{E}\left(h_{t+\varepsilon,\varepsilon}\left(X\right)\right). \tag{1.35}$$

Moreover,

$$\mathbb{E}\left(h_{t+\varepsilon,\varepsilon}(Y)\right)-\mathbb{E}\left(h_{t-\varepsilon,\varepsilon}(Y)\right)=\int_{t-2\varepsilon}^{t+2\varepsilon}\underbrace{\left(h_{t+\varepsilon,\varepsilon}(y)-h_{t-\varepsilon,\varepsilon}(y)\right)}_{\leq 1}\frac{e^{\frac{-y^2}{2}}}{\sqrt{2\pi}}\mathrm{d}y\leq\frac{4\varepsilon}{\sqrt{2\pi}}.$$

Hence, we get that

$$\mathbb{E}\left(h_{t+\varepsilon,\varepsilon}(Y)\right)-\frac{4\varepsilon}{\sqrt{2\pi}}\leq\mathbb{P}\left(Y\leq t\right)\leq\mathbb{E}\left(h_{t-\varepsilon,\varepsilon}(Y)\right)+\frac{4\varepsilon}{\sqrt{2\pi}}. \tag{1.36}$$

Then, we subtract (1.36) from (1.35) and we take the supremum over $t\in\mathbb{R}$, observing that

$$\sup_{t\in\mathbb{R}}\left|\mathbb{E}\Big(h_{t-\varepsilon,\varepsilon}\left(X\right)-h_{t-\varepsilon,\varepsilon}(Y)\Big)\right|=\sup_{t\in\mathbb{R}}\left|\mathbb{E}\Big(h_{t+\varepsilon,\varepsilon}\left(X\right)-h_{t+\varepsilon,\varepsilon}(Y)\Big)\right|.$$

We get (1.34). □

Chapitre 2

On the variations of the sum of digits in the Zeckendorf representation : an algorithm to compute the distribution and mixing properties

2.1 Introduction

Throughout this article, we denote by $\mathbb{N} := \{0, 1, 2, \cdots\}$ the set of *natural integers* and $\varphi := \frac{1+\sqrt{5}}{2}$ represents the *golden ratio*. We also define the well-known *Fibonacci sequence* as follows:

$$\begin{cases} F_1 & := 1, \\ F_2 & := 1, \\ F_k & := F_{k-1} + F_{k-2} \quad \text{if } k \geq 3. \end{cases} \tag{2.1}$$

By Zeckendorf's theorem (see [30], Theorem I.a, page 179 or [18] Theorem A page 1), every natural integer can be uniquely written as a sum of non-consecutive Fibonacci terms. That is: for any $n \in \mathbb{N}$, we consider the associated sequence of *digits* $(n_k)_{k\geq 2} \in \{0, 1\}^\infty$ without two consecutive 1's, finitely many of them being strictly positive and such that

$$n = \sum_{k\geq 2} n_k F_k. \tag{2.2}$$

For $n \neq 0$ and $\ell := \max\{k : \ n_k \neq 0\}$, we introduce the notation $\overline{n_\ell \cdots n_2} := n$ which generalises the usual way we write numbers in integer base, and which we refer to as the *(Zeckendorf) expansion* of n. This way to expand numbers is actually a particular case of an *Ostrowski's numeration system* (see [20], page 1 or [2], page 211 or [1]). By convention, we agree $\overline{0} := 0$. Then, we define the (Zeckendorf-)*sum-of-digits* function as

$$s(n) := \sum_{k\geq 2} n_k.$$

A central object in our paper is the *variation of the sum of digits* when we add a fixed integer r to n: for $r, n \in \mathbb{N}$, we set

$$\Delta^{(r)}(n) := s(n+r) - s(n). \tag{2.3}$$

The analogous variation in integer base has been studied a lot: Bésineau has shown some statistical properties (in [3]), its variance has been studied, in binary base, by Emme and Prikhod'ko (in [12]) or Spiegelhofer and Wallner (in [28]), Emme and Hubert have proved a CLT (in binary, in [11]) which has been improved by Spiegelhofer and Wallner (in binary, in [28]) or by the author in collaboration with de la Rue and Janvresse (in arbitrary integer base, in [6]). We also refer to [19, 26, 8, 27] for connected results. Some theorems have been proved for the Zeckendorf expansion of an integer: for instance, Griffiths (in [14]) about the digit proportions; or Labbé and Lepšovà about the addition in this numeration system (in [17]). Drmota, Müllner and Spiegelhofer have obtained results about the existence of primes numbers with a fixed Zeckendorf sum-of-digits (in [9]). However, not much has been done about this variation in the Zeckendorf representation except the work from Spiegelhofer (in [26], Lemma 1.30) which proves that, for any $d \in \mathbb{Z}$, the following asymptotic density exists

$$\mu^{(r)}(d) := \lim_{N\to+\infty} \frac{1}{N} \left| \left\{ n < N : \ \Delta^{(r)}(n) = d \right\} \right|.$$

In the present paper, we adapt the ergodic theory point of view introduced in [6] to the Zeckendorf expansion, recovering Spiegelhofer's result and getting new results about $\mu^{(r)}$ and $\Delta^{(r)}$. In particular, we provide an algorithm that computes $\mu^{(r)}(d)$ and we represent $\Delta^{(r)}$ as the sum of a stochastic mixing process.

To do so, following the path started in [6], we study the variations of the sum of digits function in an appropriate probability space given by the compact set $\mathbb{X}$ of (Zeckendorf-)*adic numbers*. We consider on $\mathbb{X}$ the action of the odometer, and endow $\mathbb{X}$ with its unique invariant probability measure $\mathbb{P}$.

We extend $\Delta^{(r)}$ almost everywhere on $\mathbb{X}$ and show in Section 2.4 (Proposition 2.4.1) that, for every $d \in \mathbb{Z}$

$$\mu^{(r)}(d) = \mathbb{P}\left(\left\{ x \in \mathbb{X} \ : \ \Delta^{(r)}(x) = d \right\} \right).$$

Using the Rokhlin towers of the dynamical system, we provide an algorithm to compute $\mu^{(r)}(d)$. This algorithm and its consequences can be adapted in integer base. A first implication is the following corollary on the behaviour of the (negative) tail of the distribution.

Corollary 2.1.1. *For d small enough in $\mathbb{Z}$, we have the formula*

$$\mu^{(r)}(d-1) = \mu^{(r)}(d) \times \frac{1}{\varphi^2}.$$

Remark 4. *One can show that the analogous result in base $b \geq 2$ is the same replacing φ^2 by b.*

Another implication is the next theorem about the measure $\mu^{(F_\ell)}$.

Theorem 2.1.2. *For any $\ell \geq 3$*

$$\mu^{(F_\ell)} = \mu^{(1)}. \tag{2.4}$$

The analogous result in integer base b replaces F_ℓ by b^ℓ. It is actually a trivial result in base b. However, in the Zeckendorf decomposition this result is much less obvious, due to the particular behaviour of carry propagations that we describe in Subsection 2.2.1.

On the other side, to state our mixing result, we need to define the notion of *blocks* in the expansion of an integer r and to define a probabilistic notion of (α-)*mixing coefficients* (see the survey from Bradley [5] for others).

Definition 2.1.3. *A block in the expansion of an integer $r \in \mathbb{N}$ is defined as a maximal sequence of the pattern $\overline{10}$. (If $r_2 = 1$, we agree that a maximal sequence $\overline{r_{2\ell} \cdots r_2}$ is a block if $r_{2k} = 1$ for $k = 1, \cdots, \ell$.) We define $\rho(r)$ as the number of blocks in the expansion of r.*

$$r = \overline{\underbrace{10}_{B_2}\underbrace{10}_{B_1}0\underbrace{10}} \qquad r = \overline{\underbrace{10}_{B_3}00\underbrace{10}_{B_2}000\underbrace{10101}_{B_1}}$$

$$\rho(r) = 2 \qquad \rho(r) = 3$$

Figure 2.1: Two examples of decomposition in blocks.

Definition 2.1.4. *Let $(X_i)_{i\geq 1}$ be a (finite or infinite) sequence of random variables. The associated* α-mixing coefficients *$\alpha(k)$, $k \geq 1$, are defined by*

$$\alpha(k) := \sup_{p\geq 1} \sup_{A,B} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|$$

where the second supremum is taken over all events A and B such that

- $A \in \sigma(X_i : 1 \leq i \leq p)$ *and*
- $B \in \sigma(X_i : i \geq k + p)$.

By convention, if X_i is not defined when $i \geq k + p$ then the σ-algebra is trivial.

For an integer r, we enumerate the blocks of r from the units position and we define $X_i^{(r)}$ as the action of the i^{th} block once the previous blocks have already been taken into consideration (see Subsection 2.7.1 for more details). These actions are constructed in order to have the equality

$$\Delta^{(r)} = \sum_{i=1}^{\rho(r)} X_i^{(r)}. \tag{2.5}$$

We state the following theorem which gives an upper bound on the α-mixing coefficients that is independent of r.

Theorem 2.1.5. *The α-mixing coefficients of $(X_i^{(r)})_{i=1,\cdots,\rho(r)}$ satisfy*

$$\forall k \geq 1, \qquad \alpha(k) \leq 12\left(1 - \frac{1}{\varphi^8}\right)^{\frac{k}{6}} + \frac{1}{\varphi^{2k}}.$$

A perspective we have with this result is to find a CLT for $\Delta^{(r)}$. In [6], a similar result together with some control of the variance depending on the number of blocks was used to prove a CLT for $\mu^{(r)}$ in the integer-base case. Such a control is, here, not clear at all since, contrary the situation in integer base, we do not have inductive relations on $\mu^{(r)}$.

Roadmap

Section 2.2 is devoted to the comprehension of the additions "+1" and "$+F_k$" on the set $\mathbb{X}$ of Z-adic integers. We emphasize the important role played by the patterns $w_0 := \overline{01000}$ and $w_1 := 10010$ during such additions.

In Section 2.3, we focus on the unique ergodic measure $\mathbb{P}$ of the odometer. We show $\mathbb{P}$ satisfies some renewal properties (Proposition 2.3.2) and we estimate the ϕ-mixing coefficients for the coordinates of a Z-adic integer (Proposition 2.3.5).

Then, in Section 2.4, we place the study of the measures $\mu^{(r)}$ in the context of the odometer on $\mathbb{X}$. We extend $\Delta^{(r)}$ almost everywhere on $\mathbb{X}$ and we show that the convergence

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n<N} f(\Delta^{(r)}(n)) = \int_{\mathbb{X}} f(\Delta^{(r)}(x)) \mathrm{d}\mathbb{P}(x)$$

is satisfied for functions $f : \mathbb{Z} \to \mathbb{C}$ of polynomial growth (Proposition 2.4.1) and, more generally, for functions f such that $f \circ \Delta^{(r)}$ is integrable (Proposition 2.4.5). We deduce from Proposition 2.4.1 that $\mu^{(r)}(d) = \mathbb{P}(\{x \in \mathbb{X} : \Delta^{(r)}(x) = d\})$ and that $\mu^{(r)}$ has finite moments. In particular, we show that $\mu^{(r)}$ is of zero-mean.

In Section 2.5, using several Lemmas (2.5.3,2.5.4,2.5.5) and Corollary 2.5.6, we construct an algorithm that computes $\mu^{(r)}$. A pseudo-code is given in Subsection 2.5.2. Also, in Subsection 2.5.4, we prove Corollary 2.1.1 which gives a control on the tail of $\mu^{(r)}$.

Section 2.6 is devoted to the proofs (there are two) of Theorem 2.1.2. The first proof consists in applying the algorithm while the second uses also the renewal properties we developed in Section 2.3.

In the last section, we build a finite sequence of random variables associated to the addition of some integer r that will give the decomposition of $\Delta^{(r)}$ mentioned at (2.5). Using that sequence, we prove Theorem 2.1.5 on the estimation of the α-mixing coefficients for this sequence.

Acknowledgments

The author wishes to thank T. de la Rue and E. Janvresse who introduced him to this subject. They gave the author multiple good advices.

2.2 How to do additions

2.2.1 How to add integers

Here we describe the algorithm of the addition using the Zeckendorf way to represent numbers. We start with the addition of 1. There are two cases.

1. Either there exists $\ell \geq 0$ such that the Zeckendorf decomposition of n is

$$n = \sum_{k=1}^{\ell} F_{2k+1} + \sum_{k\geq 2\ell+4} n_k F_k$$

Then, with the relations (2.1), we get

$$\begin{array}{ccccccccc} (n) & & \cdots & n_{2\ell+4} & 0 & 0 & (10)^\ell \\ & + & & & & & 1 \\ \hline (n+1) & = & \cdots & n_{2\ell+4} & 0 & 1 & (00)^\ell \end{array}$$

Indeed, $F_3 + \cdots F_{2\ell+1} + F_2 = F_{2\ell+2}$.

2. Or there exists $\ell \geq 0$ such that

$$n = \sum_{k=1}^{\ell+1} F_{2k} + \sum_{k \geq 2\ell+5} n_k F_k.$$

Then, for the same reasons

$$\begin{array}{cccccccc} (n) & & \cdots & n_{2\ell+5} & 0 & 0 & 1 & (01)^\ell \\ & + & & & & & & 1 \\ \hline (n+1) & = & \cdots & n_{2\ell+5} & 0 & 1 & 0 & (00)^\ell \end{array}$$

We observe that adding 1 to the rightmost digit of a block as defined in 2.1.3 modifies that block into a chain of 0 digits of the same length and put a 1 at the first left position of the block.

Of course, in order to compute the addition of two integers, adding 1 as many times as needed is enough. However, we want to show a main difference with the addition in integer base. In integer base, adding 1 at some position $k \geq 2$ to an integer n may change the digits of the expansion of n of higher indices due to a carry propagations. Here, it can also change the digits of lower indices. We consider the addition $n + F_k$ where $k \geq 3$. We observe that a consequence of (2.1) is

$$2F_k = \begin{cases} F_2 + F_4 & \text{if } k = 3, \\ F_{k+1} + F_{k-2} & \text{otherwise.} \end{cases} \tag{2.6}$$

Many cases appear. For simplicty, the digit in **color** will represent the digit at position k and we do not represent digits that remain the same in the expansion of n and $n + F_k$. We start with the cases that change only digits of indices $\geq k$.

1. If $n = \overline{\cdots 0\mathbf{0}0\cdots}$ then

$$\begin{array}{cccccccc} (n) & & \cdots & 0 & \mathbf{0} & 0 & \cdots \\ (+F_k) & + & & & \mathbf{1} & & \\ \hline (n+F_k) & = & \cdots & 0 & \mathbf{1} & 0 & \cdots \end{array}$$

2. If there exists $\ell \geq 0$ such that $n = \overline{\cdots 001(01)^\ell \mathbf{0}\cdots}$ then

$$\begin{array}{cccccccccc} (n) & & \cdots & 0 & 0 & 1 & (01)^\ell & \mathbf{0} & \cdots \\ (+F_k) & + & & & & & & \mathbf{1} & \\ \hline (n+F_k) & = & \cdots & 0 & 1 & 0 & (00)^\ell & \mathbf{0} & \cdots \end{array}$$

We continue with the first case where the digit of index $k-1$ is changed.

3. If there exists $\ell \geq 0$ such that $n = \overline{\cdots 00(10)^\ell \mathbf{0}1\cdots}$ then

$$\begin{array}{cccccccccc} (n) & & \cdots & 0 & 0 & (10)^\ell & \mathbf{0} & 1 & \cdots \\ (+F_k) & + & & & & & \mathbf{1} & & \\ \hline (n+F_k) & = & \cdots & 0 & 1 & (00)^\ell & \mathbf{0} & 0 & \cdots \end{array}$$

Now, we consider the cases where many digits of indices $< k$ are changed. It is due to (2.6).

4. If $k \geq 5$ and there exist $\ell, \ell' \geq 0$ such that $n = \overline{\cdots 00(10)^\ell \mathbf{1} (01)^{\ell'} 000 \cdots}$ then

$$\begin{array}{cccccccccccc} (n) & & \cdots & 0 & 0 & (10)^\ell & \mathbf{1} & (01)^{\ell'} & 0 & 0 & 0 & \cdots \\ (+F_k) & + & & & & & \mathbf{1} & & & & & \\ \hline (n+F_k) & = & \cdots & 0 & 1 & (00)^\ell & \mathbf{0} & (10)^{\ell'} & 0 & 1 & 0 & \cdots \end{array}$$

5. If $k \geq 6$ and there exist $\ell, \ell' \geq 0$ such that $n = \overline{\cdots 00(10)^\ell \mathbf{1} (01)^{\ell'} 0010 \cdots}$ then

$$\begin{array}{cccccccccccc} (n) & & \cdots & 0 & 0 & (10)^\ell & \mathbf{1} & (01)^{\ell'} & 0 & 0 & 1 & 0 & \cdots \\ (+F_k) & + & & & & & \mathbf{1} & & & & & & \\ \hline (n+F_k) & = & \cdots & 0 & 1 & (00)^\ell & \mathbf{0} & (10)^{\ell'} & 1 & 0 & 0 & 0 & \cdots \end{array}$$

Finally, we consider the "boundary" cases where all the digits of indices $< k$ are changed.

6. If $k \geq 4$, k is even and there exists $\ell \geq 0$ such that $n = \overline{\cdots 00(10)^\ell \mathbf{1} (01)^{\frac{k-2}{2}}}$ then

$$\begin{array}{ccccccccc} (n) & & \cdots & 0 & 0 & (10)^\ell & \mathbf{1} & (01)^{\frac{k-2}{2}} \\ (+F_k) & + & & & & & \mathbf{1} & \\ \hline (n+F_k) & = & \cdots & 0 & 1 & (00)^\ell & \mathbf{0} & (10)^{\frac{k-2}{2}} \end{array}$$

7. If $k \geq 5$, k is odd and there exists $\ell \geq 0$ such that $n = \overline{\cdots 00(10)^\ell \mathbf{1} (01)^{\frac{k-3}{2}} 0}$ then

$$\begin{array}{ccccccccc} (n) & & \cdots & 0 & 0 & (10)^\ell & \mathbf{1} & (01)^{\frac{k-3}{2}} & 0 \\ (+F_k) & + & & & & & \mathbf{1} & & \\ \hline (n+F_k) & = & \cdots & 0 & 1 & (00)^\ell & \mathbf{0} & (10)^{\frac{k-3}{2}} & 1 \end{array}$$

2.2.2 The Z-adic integers and how to add one of them to an integer

We define the *Zeckendorf-adic integers* (or *Z-adic integers* for simplicity) as elements of

$$\mathbb{X} := \left\{ x \in \{0;1\}^{\mathbb{N}_{\geq 2}} : \forall k \geq 2 \; x_k x_{k+1} = 0 \right\}.$$

Also, coordinates of a Z-adic integer $x \in \mathbb{X}$ are interpreted as digits in the Zeckendorf representation: elements of $\mathbb{X}$ can be viewed as "generalized integers having possibly infinitely many non zero digits in Zeckendorf representation". An element $x = (x_k)_{k \geq 2} \in \mathbb{X}$ will be represented as a left-infinite sequence $(\cdots, x_3, x_2)$, x_2 being the unit digit. We endow $\mathbb{X}$ with the product topology which turns it into a compact metrizable space. The set $\mathbb{N}$ can be identified with the subset of sequences with finite support. More precisely, using the inclusion function

$$i : n = \overline{n_\ell \cdots n_2} \in \mathbb{N} \longmapsto (\cdots, 0, n_\ell, \cdots, n_2) \in \mathbb{X}$$

we identify $\mathbb{N}$ and $i(\mathbb{N})$. We can also identify the notation

$$(\cdots, x_3, x_2) = \overline{\cdots x_3 x_2}.$$

We take the opportunity to define $\mathbb{X}_f$ as the finite sequences of 0's and 1's without two consecutive 1's. For instance, the Zeckendorf expansion of a given integer is composed using a sequence in $\mathbb{X}_f$.

Let us define, for $\ell \geq 2$ and $(n_k)_{k\geq 2} \in \mathbb{X}_f$, the *cylinder* $C_{n_\ell \cdots n_2}$ as the set of sequences $x \in \mathbb{X}$ such that $x_i = n_i$ for $i = 2, \cdots, \ell$. We observe that $n_\ell \cdots n_2$ is not necessarily the Zeckendorf expansion of a given integer : the leftmost digit(s) can be 0('s).

We want to extend on $\mathbb{X}$ the transformation $n \mapsto n+1$ defined on $\mathbb{N}$. Thanks to the description given in Subsection 2.2.1, it is convenient to consider the transformation T on $\mathbb{X}$ defined by the following formula, where $\ell \in \mathbb{N}$

$$\begin{array}{rccc} T: & \mathbb{X} & \longrightarrow & \mathbb{X} \\ & \overline{\cdots x_{2\ell+4}00(10)^\ell} & \longmapsto & \overline{\cdots x_{2\ell+4}01(00)^\ell} \\ & \overline{\cdots x_{2\ell'+3}001(01)^\ell} & \longmapsto & \overline{\cdots x_{2\ell'+3}010(00)^\ell} \\ & \overline{(10)^\infty} & \longmapsto & \overline{0^\infty} \\ & \overline{(01)^\infty} & \longmapsto & \overline{0^\infty} \end{array}$$

Indeed, thanks to the Subsection 2.2.1, we observe that $T_{|\mathbb{N}}(n) = n+1$. Now, if we take $x \in \mathbb{X}$ whose expansion contains two consecutive 0's, we observe that the sequence $(\overline{x_\ell \cdots x_2} + 1)_{\ell \geq 2}$ converges to $T(x)$ because the digits will not changed eventually so we can define $x + 1 := T(x)$ in that case. Otherwise, if $x \in \mathbb{X}$ does not have two consecutive 0's in its expansion, there are two cases : $\overline{(10)^\infty}$ and $\overline{(01)^\infty}$. Adding 1 to the truncated sequence $\left(\overline{(10)^\ell}\right)_{\ell\in\mathbb{N}}$ converges to $\overline{(10)^\infty} + 1 := \overline{0^\infty}$. It is the same for $\overline{(01)^\infty}$. Thus, the transformation T can be described in a simpler way as

$$T : \left\{ \begin{array}{ccc} \mathbb{X} & \longrightarrow & \mathbb{X} \\ x & \longmapsto & x+1 \end{array} \right.$$

Due to the two pre-images of 0^∞, T is not a homeomorphism on $\mathbb{X}$ but we have the following result that ensures that $(\mathbb{X}, T)$ is a topological dynamical system that we call the *Odometer*.

Proposition 2.2.1. *T is continuous on $\mathbb{X}$, surjective on $\mathbb{X}$ and one-to-one on $\mathbb{X}\backslash\{0^\infty\}$.*

Proof. The surjectivity on $\mathbb{X}$ and the injectivity on $\mathbb{X}\backslash\{0^\infty\}$ follow immediately from the definition of T. Observing that the cylinders generate the topology, another consequence of the definition of T is that the pre-image of a cylinder is another cylinder. So T is also continuous on $\mathbb{X}$. □

We know how to add 1 to a Z-adic integer x. Repeating this operation enables us to add an integer r to x. For later purposes, we need to specify how to add F_k (≥ 3) directly. In Subsection 2.2.1, we have compute $x + F_k$ for the $x \in \mathbb{X}$ which has two consecutive 0's at indices $> k$. Thus, we only need to focus on what happens if x does not have two consecutive 0's at indices $> k$. There is only a finite number of cases to consider which we detail below. For simplicity again, we write in **color** the digits at position k and we do not represent digits that are not modified. We start with the cases where the only digits that change are those of indices $\geq k-1$.

1. If $x = \overline{(10)^\infty \mathbf{0}1\cdots}$ then

$$\begin{array}{ccccccc} (x) & & (10)^\infty & \mathbf{0} & 1 & \cdots \\ (F_k) & + & & \mathbf{1} & & \\ \hline (x+F_k) & = & (00)^\infty & \mathbf{0} & 0 & \cdots \end{array}$$

2. If $x = \overline{(01)^\infty \mathbf{0}\cdots}$ then

$$\begin{array}{ccccccc} (x) & & (01)^\infty & \mathbf{0} & \cdots \\ (F_k) & + & & \mathbf{1} & \\ \hline (x+F_k) & = & (00)^\infty & \mathbf{0} & \cdots \end{array}$$

Now we consider cases where digits of small indices are changed but not all them.

3. If $k \geq 5$ and there exists $\ell' \geq 0$ such that $x = \overline{(10)^\infty \mathbf{1}(01)^{\ell'}000\cdots}$ then

$$\begin{array}{cccccccccc} (x) & & (10)^\infty & \mathbf{1} & (01)^{\ell'} & 0 & 0 & 0 & \cdots \\ (F_k) & + & & \mathbf{1} & & & & & \\ \hline (x+F_k) & = & (00)^\infty & \mathbf{0} & (10)^{\ell'} & 0 & 1 & 0 & \cdots \end{array}$$

4. If $k \geq 6$ and there exists $\ell' \geq 0$ such that $x = \overline{(10)^\infty \mathbf{1}(01)^{\ell'}0010\cdots}$ then

$$\begin{array}{cccccccccc} (x) & & (10)^\infty & \mathbf{1} & (01)^{\ell'} & 0 & 0 & 1 & 0 & \cdots \\ (F_k) & + & & \mathbf{1} & & & & & & \\ \hline (x+F_k) & = & (00)^\infty & \mathbf{0} & (10)^{\ell'} & 1 & 0 & 0 & 0 & \cdots \end{array}$$

Finally, we consider cases where the whole prefix of x is modified.

5. If $k \geq 4$ and is even and $x = \overline{(10)^\infty \mathbf{1}(01)^{\frac{k-2}{2}}}$ then

$$\begin{array}{ccccc} (x) & & (10)^\infty & \mathbf{1} & (01)^{\frac{k-2}{2}} \\ (F_k) & + & & \mathbf{1} & \\ \hline (x+F_k) & = & (00)^\infty & \mathbf{0} & (10)^{\frac{k-2}{2}} \end{array}$$

6. If $k \geq 3$ and is odd and $x = \overline{(10)^\infty \mathbf{1}(01)^{\frac{k-3}{2}}0}$ then

$$\begin{array}{ccccccc} (x) & & (10)^\infty & \mathbf{1} & (01)^{\frac{k-3}{2}} & 0 \\ (F_k) & + & & \mathbf{1} & & \\ \hline (x+F_k) & = & (00)^\infty & \mathbf{0} & (10)^{\frac{k-3}{2}} & 1 \end{array}$$

We can sum up all these cases in the following proposition. We need to take $k \geq 6$ to ensure each case appear.

Proposition 2.2.2. *The table below summarises the action of the addition of F_k ($k \geq 2$) on the digits of $x \in \mathbb{X}$. Here $\ell, \ell' \in \mathbb{N}$ and we represent in **color**, the digits at position k. The digits that*

are not explicitly written remain untouched by the operation.

$$
\begin{array}{lll}
T^{F_k}: \mathbb{X} & \longrightarrow & \mathbb{X} \\
\overline{\cdots 0\mathbf{0}0\cdots} & \longmapsto & \overline{\cdots 0\mathbf{1}0\cdots} \quad (2.7)\\
\overline{\cdots 001(01)^{\ell}\mathbf{0}\cdots} & \longmapsto & \overline{\cdots 010(00)^{\ell}\mathbf{0}\cdots} \quad (2.8)\\
\overline{\cdots 00(10)^{\ell}\mathbf{0}1\cdots} & \longmapsto & \overline{\cdots 01(00)^{\ell}\mathbf{0}0\cdots} \quad (2.9)\\
\overline{\cdots 00(10)^{\ell}\mathbf{1}(01)^{\ell'}000\cdots} & \longmapsto & \overline{\cdots 01(00)^{\ell}\mathbf{0}(10)^{\ell'}010\cdots} \quad (2.10)\\
\overline{\cdots 00(10)^{\ell}\mathbf{1}(01)^{\ell'}0010\cdots} & \longmapsto & \overline{\cdots 01(00)^{\ell}\mathbf{0}(10)^{\ell'}1000\cdots} \quad (2.11)\\
\overline{\cdots 00(10)^{\ell}\mathbf{1}(01)^{\frac{k-5}{2}}001} & \longmapsto & \overline{\cdots 01(00)^{\ell}\mathbf{0}(10)^{\frac{k-5}{2}}100} \quad (2.12)\\
\overline{\cdots 00(10)^{\ell}\mathbf{1}(01)^{\frac{k-4}{2}}00} & \longmapsto & \overline{\cdots 01(00)^{\ell}\mathbf{0}(10)^{\frac{k-4}{2}}01} \quad (2.13)\\
\overline{\cdots 00(10)^{\ell}\mathbf{1}(01)^{\frac{k-2}{2}}} & \longmapsto & \overline{\cdots 01(00)^{\ell}\mathbf{0}(10)^{\frac{k-2}{2}}} \quad (2.14)\\
\overline{\cdots 00(10)^{\ell}\mathbf{1}(01)^{\frac{k-3}{2}}0} & \longmapsto & \overline{\cdots 01(00)^{\ell}\mathbf{0}(10)^{\frac{k-3}{2}}1} \quad (2.15)\\
\overline{(01)^{\infty}\mathbf{0}\cdots} & \longmapsto & \overline{0^{\infty}\mathbf{0}\cdots} \quad (7.bis)\\
\overline{(10)^{\infty}\mathbf{0}1\cdots} & \longmapsto & \overline{0^{\infty}\mathbf{0}0\cdots} \quad (8.bis)\\
\overline{(10)^{\infty}\mathbf{1}(01)^{\ell'}000\cdots} & \longmapsto & \overline{0^{\infty}\mathbf{0}(10)^{\ell'}010\cdots} \quad (9.bis)\\
\overline{(10)^{\infty}\mathbf{1}(01)^{\ell'}0010\cdots} & \longmapsto & \overline{0^{\infty}\mathbf{0}(10)^{\ell'}1000\cdots} \quad (10.bis)\\
\overline{(10)^{\infty}\mathbf{1}(01)^{\frac{k-5}{2}}001} & \longmapsto & \overline{0^{\infty}\mathbf{0}(10)^{\frac{k-5}{2}}100} \quad (11.bis)\\
\overline{(10)^{\infty}\mathbf{1}(01)^{\frac{k-4}{2}}00} & \longmapsto & \overline{0^{\infty}\mathbf{0}(10)^{\frac{k-4}{2}}01} \quad (12.bis)\\
\overline{(10)^{\infty}\mathbf{1}(01)^{\frac{k-2}{2}}} & \longmapsto & \overline{0^{\infty}\mathbf{0}(10)^{\frac{k-2}{2}}} \quad (13.bis)\\
\overline{(10)^{\infty}\mathbf{1}(01)^{\frac{k-3}{2}}0} & \longmapsto & \overline{0^{\infty}\mathbf{0}(10)^{\frac{k-3}{2}}1} \quad (14.bis)
\end{array}
$$

(We precise that k is even in cases (2.13)*,* (2.14)*,* (12.bis) *and* (13.bis) *and k is odd in cases* (2.12)*,* (2.15)*,* (11.bis) *and* (14.bis)*.)*

2.2.3 Stopping conditions when adding an integer to an adic number

Through the cases described in Proposition 2.2.2, we observe that if there is a 1 at position k in the expansion of x, the addition of F_k yields a carry propagation in both directions:

- to the left, modifying digits of higher indices (as in integer base)
- to the right, modifying digits of lower indices.

The propagation (in both directions) happens through a maximal sequence of alternative 0's and 1's and is stopped at the first occurrence of two consecutive 0's. But the modifications depend on the propagation direction.

- In the propagation to the left, the maximal subword of alternative 1's and 0's will be transformed into a subword of 0's of the same length (case (2.10) for instance), and the stopping pattern 00 is transformed into 01. We precise this propagation also happens if $x_k = 0$ (case (2.8) for instance).

- In the propagation to the right, which only happens if $x_k = 1$, the maximal subword of alternative 1's and 0's is transformed into a symmetrical subword where the 1's become 0's and vice-versa (cases (2.10) and (2.11) for instance). Then the first occurrence of 00 (in the sense the largest index $\leq k$ such that the digits of x are 00) can either be part of the pattern $w_0 := \overline{01000}$ or $w_1 := \overline{10010}$. (We call w_0 and w_1 the *right-stopping pattern.*)

Depending on the right-stopping pattern ($\overline{01000}$ or $\overline{10010}$), the modifications of digits at these indices are given by the next scheme.

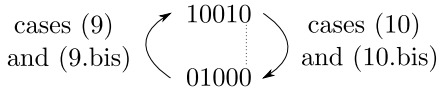


Figure 2.2: Modifications at the position of the right-stopping pattern.

Note that in both cases, we get a new right-stopping pattern at the same position as before the addition of F_k. This addition of F_k to $x \in \mathbb{X}$ modifies some digits at positions $\leq k-2$ only if $x_k = 1$, and, in this case, the modification of the digits takes place up to the first occurrence of one of the right-blocking patterns.

Formally: let $x \in \mathbb{X}$ and $k \geq 2$.

- If $x_k = 0$ then $(x + F_k)_n = x_n$ for all $n \leq k-2$.
- If $x_k = 1$ and if there exists $j \leq k+1$ such that $x_j x_{j-1} \cdots x_{j-4}$ is a right-stopping patterns w_i ($i \in \{0,1\}$), -we denote by j' the largest index with this property- then
  - $(x + F_k)_n = x_n$ for all $n \leq j' - 4$
  - w_{1-i} appears in $(x + F_k)$ at the same position j', unless $w_i = w_0$ and $j' = k+1$, in which case we might have 00010 instead of w_1 at position j' in $x + F_k$ (case (2.10) with $\ell = \ell' = 0$).

A straightforward consequence is the following Proposition.

Proposition 2.2.3. *Assume that the right-stopping pattern w_i ($i = 0, 1$) appears in $x \in \mathbb{X}$ at position $j \geq 5$, that is: $x_j x_{j-1} \cdots x_{j-4} = w_i$. Let $k \geq j-1$. Then*

- *for all $n \leq j-4$, $(x + F_k)_n = x_n$ and*
- *w_0, w_1 or 00010 appears in $x + F_k$ at some position j' with $k+1 \geq j' \geq j$.*

We now state the following corollary that enhances that property when we add not only a Fibonacci term but an integer whose expansion involves Fibonacci numbers of high indices.

Corollary 2.2.4. *Let $x \in \mathbb{X}$. Assume that, for some $\ell \geq 2$, $x_\ell = x_{\ell+1} = 0$. Let $r \in \mathbb{N}$ be such that $r_j = 0$ for $j = 2, \cdots, \ell+1$. Then, for each $n \leq \ell - 2$, we have $(x + r)_n = x_n$.*

Proof. We are considering the following addition:

(x)		$\cdots$	$x_{\ell+3}$	$x_{\ell+2}$	0	0	$x_{\ell-1}$	$x_{\ell-2}$	$x_{\ell-3}$	$\cdots$
(r)	$+$	$\cdots$	$r_{\ell+3}$	$r_{\ell+2}$	0	0	0	0	0	$\cdots$
$(x+r)$	$=$	$\cdots$	$\star$	$\star$	$\star$	$\star$	$\star$	$x_{\ell-2}$	$x_{\ell-3}$	$\cdots$

Let $r = F_{k_{s(r)}} + \cdots + F_{k_1}$ be the Zeckendorf decomposition of r with $k_{s(r)} > \cdots > k_2 > k_1 \geq \ell + 2$. We first consider the addition of F_{k_1} to x:

- if, for all j such that $k_1 \geq j \geq \ell$, we have $x_j = 0$, then we have for every $n \leq \ell + 1$, $(x + F_{k_1})_n = x_n$ (in particular 00 appears at the same place in $x + F_{k_1}$)
- otherwise, there exists a largest integer j' with $k_1 + 1 \geq j' \geq \ell + 2$, such that one of the right-stopping pattern appears in x at position j'. We can then apply the above proposition which proves that
  - either w_0 or w_1 appears at position j' in $x + F_{k_1}$
  - or 00 appears at position j' in $x + F_{k_1}$, and $j' \geq \ell + 3$

  In each case, we still have $(x + F_{k_1})_n = x_n$ for $n \leq \ell - 2$.

Then we prove by induction on t such that for each t, $1 \leq t \leq s(r)$, the above is true for $x + F_{k_1} + \cdots + F_{k_t}$. □

The following lemma ensures that the pattern 00 is a left-stopping condition: it stops the propagation of a carry coming from the right.

Lemma 2.2.5. *Let $x \in \mathbb{X}$, $r \in \mathbb{N}$. Let $\ell \geq 2$ be such that $r < F_{\ell+1}$ and assume $x_{\ell+2} = x_{\ell+3} = 0$. Then, for all $k \geq \ell + 3$, we have $(x + r)_k = x_k$.*

Proof. The assumption $r < F_{\ell+1}$, implies $r = \overline{r_\ell \cdots r_2}$ with $(r_\ell, \cdots, r_2) \in \mathbb{X}_f$. Since $\overline{x_{\ell+1} \cdots x_2} + r < F_{\ell+2} + F_{\ell+1} = F_{\ell+3}$, a carry cannot propagate on digits of indices $\geq \ell + 3$: we have the addition

(x)		$\cdots$	$x_{\ell+4}$	0	0	$x_{\ell+1}$	x_ℓ	$\cdots$	x_2
(r)	+						r_ℓ	$\cdots$	r_2
$(x+r)$	=	$\cdots$	$x_{\ell+4}$	0	$(x+r)_{\ell+2}$	$(x+r)_{\ell+1}$	$(x+r)_\ell$	$\cdots$	$(x+r)_2$

□

The next lemma enhances the previous one: it shows that, given x with some restrictions, right-stopping patterns can appear in the expansion $x + F_k$ when k is a small integer.

Lemma 2.2.6. *Let $x \in \mathbb{X}$ and $r \in \mathbb{N}$. Let $\ell \geq 2$ be such that $x_{\ell+1} = x_{\ell+2} = x_{\ell+3} = x_{\ell+4} = 0$ and $r_{\ell+2} = 1$. We also suppose $r < F_{\ell+3}$. Then, for $k = 1, 2, 3, 4$*

$$(x + r)_{\ell+k} = r_{\ell+k}$$

or

$$(x + r)_{\ell+1} = 0 \text{ and } (x + r)_{\ell+3} = 1.$$

Proof. We are actually considering the following addition

(x)		$\cdots$	$x_{\ell+5}$	0	0	0	0	x_ℓ	$\cdots$	x_2
(r)	+	$\cdots$				1	0	r_ℓ	$\cdots$	r_2

We want to show that the expansion of $x + r$ is either

$$\overline{\cdots x_{\ell+5} 0010 (x+r)_\ell \cdots (x+r)_2} \tag{C1}$$

or

$$\overline{\cdots x_{\ell+5}0100(x+r)_\ell \cdots (x+r)_2}. \tag{C2}$$

Thanks to Lemma 2.2.5, we have $(x+r)_k = x_k$ for every $k \geq \ell+4$ which means that the digits of $x+r$ of indices $\leq \ell+3$ are given by the addition $r + \overline{x_\ell \cdots x_2}$. Now, we claim

$$(x+r)_{\ell+3} \text{ or } (x+r)_{\ell+2} \text{ is 1.}$$

Indeed, if it is not the case then $\overline{x_\ell \cdots x_2} + r < F_{\ell+2}$ while $r \geq F_{\ell+2}$ (since $r_{\ell+2} = 1$). We consider now both possibilities.

- If $(x+r)_{\ell+2} = 1$ then we obtain (C1).
- If $(x+r)_{\ell+3} = 1$ then we must have $(x+r)_{\ell+1} = 0$ since we have

$$\begin{aligned} \overline{x_\ell \cdots x_2} + r - F_{\ell+3} &= \overline{x_\ell \cdots x_2} + \overline{r_\ell \cdots r_2} + F_{\ell+2} - F_{\ell+3} \\ &< F_{\ell+1}. \end{aligned}$$

  Thus, the expansion of $x+r$ is (C2).

□

We deduce the next corollary.

Corollary 2.2.7. *Let $x \in \mathbb{X}$ and $r \in \mathbb{N}$ such that it exists $\ell \geq 2$ with $x_\ell = \cdots = x_{\ell+5} = 0$, $r_{\ell+1} = r_{\ell+5} = 0$ and $r_{\ell+3} = 1$. Denote $\widetilde{r} := \overline{r_\ell \cdots r_2}$. Then, for any $k \geq \ell+4$*

$$(x + \widetilde{r} + F_{\ell+3})_k = (x+\widetilde{r})_k = x_k \tag{2.16}$$

and, for any $k \leq \ell+2$

$$(x+r)_k = (x + \widetilde{r} + F_{\ell+3})_k = (x+\widetilde{r})_k. \tag{2.17}$$

Remark 5. *The hypothesis means that we are considering the following addition*

$$\begin{array}{cccccccccccccc} (x) & & \cdots & x_{\ell+6} & 0 & 0 & 0 & 0 & 0 & 0 & x_{\ell-1} & \cdots & x_2 \\ (r) & + & \cdots & r_{\ell+6} & 0 & 0 & 1 & 0 & 0 & r_\ell & \cdots & \cdots & r_2 \end{array}$$

and the conclusion ensures that the digits of indices $\leq \ell+2$ (i.e. those on the right-hand side of the pattern we have imposed) of $x+r$ are the same if we compute this previous addition as if we compute

$$\begin{array}{cccccccccccccc} (x) & & \cdots & x_{\ell+6} & 0 & 0 & 0 & 0 & 0 & 0 & x_{\ell-1} & \cdots & x_2 \\ (\widetilde{r} + F_{\ell+3}) & + & & & & & 1 & 0 & 0 & r_\ell & \cdots & \cdots & r_2 \end{array}$$

or if we compute

$$\begin{array}{cccccccccccccc} (x) & & \cdots & x_{\ell+6} & 0 & 0 & 0 & 0 & 0 & 0 & x_{\ell-1} & \cdots & x_2 \\ (\widetilde{r}) & + & & & & & & & & r_\ell & \cdots & \cdots & r_2 \end{array}$$

In other words, the corollary states that the conditions we put on x and r stop the propagation of carries in both direction.

Proof. First, we have the following addition

$$\begin{array}{cccccccccccccc}
(x) & & \cdots & x_{\ell+6} & 0 & 0 & 0 & 0 & 0 & 0 & x_{\ell-1} & \cdots & x_2 \\
(\widetilde{r}) & + & & & & & & & & r_\ell & \cdots & \cdots & r_2 \\
\hline
(x+\widetilde{r}) & = & \cdots & x_{\ell+6} & 0 & 0 & 0 & 0 & (x+\widetilde{r})_{\ell+1} & \cdots & \cdots & \cdots & (x+\widetilde{r})_2
\end{array}$$

Indeed, due to Lemma 2.2.5, we have that $(x+\widetilde{r})_j = x_j$ for every $j \geq \ell+2$. Then, we add $F_{\ell+3}$, it gives $x+\widetilde{r}+F_{\ell+3}$ whose Zeckendorf expansion is

$$(x+\widetilde{r}+F_{\ell+3}) = \overline{\cdots x_{\ell+6}0010(x+\widetilde{r})_{\ell+1}\cdots(x+\widetilde{r})_2}.$$

We thus get the relation (2.16).

We now consider the addition (denoting $\widetilde{x} := x+\widetilde{r}+F_{\ell+3}$)

$$\begin{array}{cccccccccc}
(\widetilde{x}) & & \cdots & x_{\ell+6} & 0 & 0 & 1 & 0 & (x+\widetilde{r})_{\ell+1} & \cdots & (x+\widetilde{r})_2 \\
& + & \cdots & r_{\ell+6} & 0 & 0 & 0 & 0 & 0 & \cdots & 0
\end{array}$$

Now, using Corollary 2.2.4, we obtain (2.17). □

The statement of Corollary 2.2.7 means that, given a given block of length 1 (in the sense that it has one pattern 10) in r, we are able to control the propagation of carries so that the left part of the addition does not change the expansion on the right part and vice versa. We now want to have the same kind of control for larger blocks. We could assume that x has many 0's facing the block in r that we want to control. However, this condition would more and more "expensive" (in the sense that the probability for x to satisfy it would decrease to 0) as the length of the block increases. To avoid this, we are looking for conditions on x that affect only a bounded number of digits, regardless of the length of the block. In other words, we want our conditions to appear in "most" of $x \in \mathbb{X}$. We obtain the following corollary where the length of the block is $m+2$ and where we fixed 8 digits in x.

Corollary 2.2.8. *Let $x \in \mathbb{X}$, $r \in \mathbb{N}$ and $m \geq 0$. Suppose that it exists $\ell \geq 2$ such that*

- *$x_i = 0$ where $i \in \{\ell, \ell+1, \ell+2, \ell+3, \ell+2m+4, \ell+2m+5, \ell+2m+6, \ell+2m+7\}$,*
- *$r_{\ell+1} = r_{\ell+2m+7} = 0$ and $r_{\ell+2i+3} = 1$ for $i = 0, \cdots, m+1$.*

Denote $\widetilde{r} := \overline{r_\ell \cdots r_2}$. Then we have, for all $k \geq \ell+2m+7$

$$(x+\widetilde{r}+\sum_{i=0}^{m+1} F_{\ell+2i+3})_k = (x+\widetilde{r})_k = x_k \tag{2.18}$$

and, for all $k \leq \ell+2m+2$

$$(x+r)_k = (x+\widetilde{r}+\sum_{i=0}^{m+1} F_{\ell+2i+3})_k. \tag{2.19}$$

Proof. For simplicity, we write B as the block $\sum_{i=0}^{m+1} F_{\ell+2i+3}$. We decompose the addition $x+r$ in several steps. First we add $\widetilde{r}$. With the hypothesis on x, we actually consider the following addition

$$\begin{array}{ccccccccccccccccc}
(x) & & \cdots & x_{\ell+8+2m} & 0 & 0 & 0 & 0 & x_{\ell+2m+3} & \cdots & x_{\ell+4} & 0 & 0 & 0 & 0 & x_{\ell-1} & \cdots & x_2 \\
(\widetilde{r}) & + & & & & & & & & & & & & & r_\ell & r_{\ell-1} & \cdots & r_2
\end{array}$$

Thanks to Lemma 2.2.5, we know that this addition can only modify digits of x of indices $\leq \ell+1$. We continue with the addition of B. For simplicity, we write $\widetilde{x} = x + \widetilde{r}$

$$\begin{array}{cccccccccccccccccc}
(\widetilde{x}) & & \cdots & x_{\ell+8+2m} & 0 & 0 & 0 & 0 & x_{\ell+2m+3} & \cdots & x_{\ell+4} & 0 & 0 & \widetilde{x}_{\ell+1} & \widetilde{x}_{\ell} & \widetilde{x}_{\ell-1} & \cdots & \widetilde{x}_2 \\
(B) & + & & & & & 1 & 0 & 1 & \cdots & 0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0
\end{array}$$

Now, Lemma 2.2.6 ensures that the expansion of $\widetilde{x}$ is not modified for digits of indices $\geq \ell+2m+7$. We thus prove (2.18). But more precisely, Lemma 2.2.6 concludes that the expansion of $\widetilde{x}+B$ is either

$$\overline{\cdots x_{\ell+8+2m}0100(\widetilde{x}+B)_{\ell+2m+3}\cdots(\widetilde{x}+B)_2}$$

or

$$\overline{\cdots x_{\ell+8+2m}0010(\widetilde{x}+B)_{\ell+2m+3}\cdots(\widetilde{x}+B)_2}.$$

In both case, a pattern 00 appears between the indices $\ell+2m+4$ and $\ell+2m+7$. Thus, we can apply Corollary 2.2.4: the final step of the addition, which consists into adding what remains in r, will not modify the digits of indices $\leq \ell+2m+2$. We obtain the conclusion (2.19). □

2.3 Unique Ergodicity of the Odometer

2.3.1 Rokhlin towers and the ergodic measure

In this Subsection, we focus on the action of T on cylinders. For each $k \geq 1$, we consider the partition of $\mathbb{X}$ into F_{k+2} cylinders corresponding to all possible blocks formed by the rightmost k digits of a Z-adic integers (we call them "*cylinders of order k*"). For example, at order 1, we partition $\mathbb{X}$ into C_0 and C_1. At order 2, we get $\mathbb{X} = C_{00} \sqcup C_{01} \sqcup C_{10}$. At order 3

$$\mathbb{X} = C_{000} \sqcup C_{001} \sqcup C_{010} \sqcup C_{100} \sqcup C_{101}.$$

In general, we order lexicographically the cylinders of order k:

- first, those whose name has a 0 at the leftmost position (there are F_{k+1} of them),
- then those whose name has a 1 at the leftmost position (there are F_k of them).

We observe that each cylinders of order k with a 0 at the leftmost position, except the last one, is mapped by T onto the next one, giving rise to a Rokhlin tower of height F_{k+1} : we call it the *large* tower of order k. Similarly, each cylinders or order k with a 1 at the leftmost position, except the last one, is mapped by T onto the next one, giving rise to another Rokhlin tower of height F_k : we call it the *small* tower of order k. Thus, for each $k \geq 1$, we get a partition of $\mathbb{X}$ into two Rokhlin towers whose levels are all cylinders of order k. For instance, for $k = 4$, we get

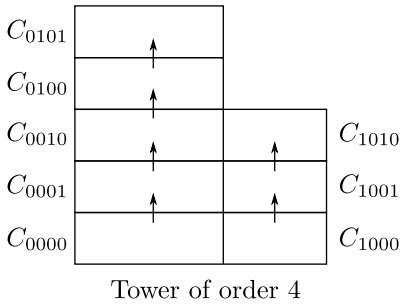


Figure 2.3: Rokhlin tower of order 4 (the action of T is representing by the arrows).

It remains to describe the transition from the Rokhlin towers of order k to those of order $k+1$. First, note that all levels of the small tower of order k are also levels of the large tower of order $k+1$, since a cylinder of order k with a 1 at the leftmost position coincides with the cylinders of order $k+1$ with an additional 0 concatenated at the left of its name (e.g. $C_{101000} = C_{0101000}$).

Each level of the large tower of order k is partitioned into two cylinders of order $k+1$: one obtained by concatenating a 0 at the left of its name and the other obtained by concatenating a 1. Thus, the large Rokhlin tower of order k is cut into two subtowers :

- the first one is the bottom part of the large Rokhlin tower of order $k+1$ (the top part being nothing but the small Rokhlin towers of order k);
- the second one is the small Rokhlin tower of order $k+1$.

This transition will be refered as the *Cut and Stack process*.

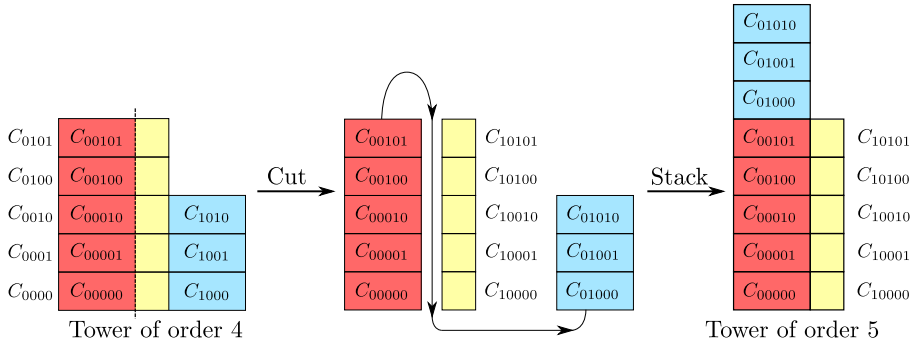


Figure 2.4: Visual description of the Cut and Stack process (for $k = 4$).

The proceding analysis of the action of T on cylinders, yielding to the construction of the Rokhlin towers, enables us to describe the unique T-invariant probability measure.

Proposition 2.3.1. $(\mathbb{X}, T)$ *is uniquely ergodic and the unique* T*-invariant measure* $\mathbb{P}$ *satisfies*

$\forall \ell \geq 2$ *and* $(r_\ell, \cdots, r_2) \in \mathbb{X}_f$

$$\mathbb{P}(C_{r_\ell \cdots r_2}) = \begin{cases} \dfrac{1}{\varphi^{\ell-1}} & \text{if } r_\ell = 0 \\ \dfrac{1}{\varphi^{\ell}} & \text{if } r_\ell = 1 \end{cases} \tag{2.20}$$

Proof. Let $\mathbb{P}$ be a T-invariant measure. Due to T-invariance, we observe that, for each order $k \geq 1$, the measure of each level of the large tower of order k is the same so as the measure of each level of the small tower of order k. We claim that

$$\mathbb{P}(C_1) = \frac{1}{\varphi^2}. \tag{2.21}$$

Indeed, if we denote by u_k (resp. v_k) the number of levels included in C_1 in the large (resp. small) tower of order k, then, due to the Cut and Stack process, we have the relations for $k \geq 2$

$$\begin{cases} u_{k+1} &= u_k + v_k \\ v_{k+1} &= u_k \end{cases}$$

with initial conditions $u_1 = 0$ and $v_0 = 1$. We deduce that, for $k \geq 1$, $u_k = F_{k-1}$ and $v_k = F_{k-2}$. Also, for all $k \geq 1$, we have the identity

$$\begin{aligned} \mathbb{P}(C_1) = u_k \ &\mathbb{P}\ (x \in \text{one level of the large tower of order k}) \\ &+ v_k\ \mathbb{P}\ (x \in \text{one level of the small tower of order k}) \\ = \frac{u_k}{F_{k+1}}\ &\mathbb{P}\ (x \in \text{the large tower of order k}) \\ &+ \frac{v_k}{F_k}\ \mathbb{P}\ (x \in \text{the small tower of order k}) \\ = \frac{v_k}{F_k} + (\frac{u_k}{F_{k+1}} - \frac{v_k}{F_k})\ &\mathbb{P}\ (x \in \text{the large tower of order k}) \end{aligned}$$

Since both quotients $\frac{u_k}{F_{k+1}}$ and $\frac{v_k}{F_k}$ converge to $\frac{1}{\varphi^2}$, we obtain (2.21). Then, we deduce that $\mathbb{P}(C_0) = 1 - \frac{1}{\varphi^2} = \frac{1}{\varphi}$. At order 2, we observe that $C_1 = C_{01}$ so, by T-invariance, $\mathbb{P}(C_{00}) = \frac{1}{\varphi^2}$. It follows $\mathbb{P}(C_{10}) = 1 - \mathbb{P}(C_{00}) - \mathbb{P}(C_{01}) = \frac{1}{\varphi^3}$. Then, suppose that, at order $k \geq 2$, each level of the large (resp. small) tower measures $\frac{1}{\varphi^k}$ (resp. $\frac{1}{\varphi^{k+1}}$). We observe that a level of the small tower of order k is, due to the Cut and Stack process, exactly a level of the large tower of order $k+1$. Thus, by T-invariance, we deduce that each level of the large tower of order $k+1$ measures $\frac{1}{\varphi^{k+1}}$. Also by T-invariance, each level of the small tower of order $k+1$ has the same measure. We denote it $p \in [0, 1]$. Since there are F_{k+2} (resp. F_{k+1}) levels in the large (resp.small) tower of order $k+1$, we have the relation

$$\frac{F_{k+2}}{\varphi^{k+1}} + pF_{k+1} = 1$$

which is equivalent to the relation

$$\varphi F_{k+2} + p\varphi^{k+2} F_{k+1} = \varphi^{k+2}$$

Combining with the classical identity $\varphi F_{k+2} + F_{k+1} = \varphi^{k+2}$, we deduce that $p = \frac{1}{\varphi^{k+2}}$. By induction, we prove (2.20). □

2.3.2 Probabilistic interpretation of the measure

In the case of the b-adic odometer ($b \geq 2$), the T-invariant measure can be interpreted as an independent choice of each digit according to the uniform law on the set of possible digits (see [6], page 7). For the Z-adic odometer, it is not as easy to describe. First, the digits do not follow the same law. Indeed, for all $k \geq 2$, $\mathbb{P}(x_k = 1) = \frac{F_{k-1}}{\varphi^k}$ and depends on k. That gives another proof about the frequency of 1's in the Zeckendorf expansion : $\mathbb{P}(x_k = 1) \xrightarrow[k\to+\infty]{} \frac{1}{\varphi^2+1}$ (see [14], [18] for other proofs). Therefore, if σ is the shift on $\mathbb{X}$, the law of x is not the same as the law of $\sigma^k(x)$, for $k \geq 1$ which means $\mathbb{P}$ is not stationary. Furthermore, the choice of a digit is not independent of the other digits because a 1 must be followed by a 0. However, the lack of stationarity and independency of $\mathbb{P}$ is compensed by the following renewal property.

Proposition 2.3.2. *Let C be a cylinder, $k \geq 2$, and $(r_k, \cdots, r_2) \in \mathbb{X}_f$. Then*

1. $\mathbb{P}\left(\sigma^k x \in C \mid x \in C_{0r_k\cdots r_2}\right) = \mathbb{P}(C)$

and

1. $\mathbb{P}\left(\sigma^{k+1} x \in C \mid x \in C_{1r_k\cdots r_2}\right) = \mathbb{P}(C)$ *(if $r_k = 0$).*

Proof. Without a loss of generality, we can suppose C is a cylinder of order k_0 with a 0 at the left side of its name for some $k_0 \geq 1$ so $\mathbb{P}(C) = \frac{1}{\varphi^{k_0}}$. Then

$$\mathbb{P}_{\left(x\in C_{0r_k\cdots r_2}\right)}\left(\sigma^{k+2}x \in C\right) = \frac{\mathbb{P}\left(x \in C_{0r_k\cdots r_2} \cap \sigma^{k+2}x \in C\right)}{\mathbb{P}(x \in C_{0r_k\cdots r_2})}.$$

We observe that the set $\{x \in C_{0r_k\cdots r_2} \cap \sigma^{k+2}x \in C\}$ is actually a cylinder of order $k + k_0$ with a leftmost 0 in its name. Its measure is therefore $\frac{1}{\varphi^{k+k_0}}$.

$$\mathbb{P}_{\left(x\in C_{0r_k\cdots r_2}\right)}\left(\sigma^{k+2}x \in C\right) = \frac{\varphi^k}{\varphi^{k+k_0}} = \frac{1}{\varphi^{k_0}} = \mathbb{P}(C).$$

We get the first point of the proposition. Then, we observe that, since $C_{1r_k\cdots r_2} = C_{01r_k\cdots r_2}$, the second is a particular case of the first one (with $k + 1$ instead of k). □

Once we know the value of x_k, the conditionnal law of x_{k+1} depends neither on k or $x_{k-1}, \cdots, x_2$. In particular, we get that

$$\mathbb{P}(x_{k+2} = 1 \mid x_{k+1}, \cdots, x_2) = \begin{cases} \frac{1}{\varphi^2} & \text{if } x_{k+1} = 0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$\mathbb{P}(x_{k+2} = 0 \mid x_{k+1}, \cdots, x_2) = \begin{cases} \frac{1}{\varphi} & \text{if } x_{k+1} = 0 \\ 1 & \text{otherwise.} \end{cases}$$

Therefore, under $\mathbb{P}$, the digits x_2, x_3, $\cdots$ form a Markov Chain with transition probabilities given on the figure, starting with the initial law

$$\mathbb{P}(x_2 = 1) = \frac{1}{\varphi^2} \text{ and } \mathbb{P}(x_2 = 0) = \frac{1}{\varphi}.$$

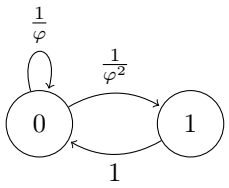


Figure 2.5: Transition probabilities of the Markov Chain.

2.3.3 Reminds on some notions of mixing coefficients

In the Introduction (more precisely at Definition 2.1.4), we introduced the notion of α-mixing coefficients. It happens that the distribution of the digits in a Z-adic integer satisfies a good inequality for another (better) notion of mixing coefficients : the ϕ-mixing coefficients. They are defined as follows.

Definition 2.3.3. *Let $(X_j)_{j\geq 1}$ be a (finite or infinite) sequence of random variables. The associated* ϕ-mixing coefficients $\phi(k)$, $k \geq 1$, *are defined by*

$$\phi(k) := \sup_{p\geq 1} \sup_{A,B} |\mathbb{P}_A(B) - \mathbb{P}(B)|$$

where the second supremum is taken over all events A and B such that

- $A \in \sigma(X_j : 1 \leq j \leq p)$,
- $\mathbb{P}(A) > 0$ *and*
- $B \in \sigma(X_j : j \geq k + p)$.

By convention, if X_j is not defined when $j \geq k + p$ then the σ-algebra is trivial.

In the case of a finite sequence $(X_1, \cdots, X_n)$, the convention implies that $\phi(k) = 0$ for $k \geq n$.

Both α and ϕ-mixing coefficients are linked together and, as written above, ϕ-mixing coefficients are "better" than α-mixing coefficients. Indeed, we have the following property from the survey [5] by Bradley.

Proposition 2.3.4 (Bradley)**.** *For every $k \geq 1$*

$$\alpha(k) \leq \frac{1}{2}\phi(k).$$

Proof. Let $p, k \geq 1$. Let $A \in \sigma(X_i \ : \ 1 \leq i \leq p)$ and $B \in \sigma(X_i \ : \ i \geq k+p)$. We claim that

$$|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \leq \frac{1}{2}\phi(k). \tag{2.22}$$

Indeed if $\mathbb{P}(A) = 0$ then $\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B) = 0$ and (2.22) is trivial. Else, by definition of $\phi(k)$, we have the inequality

$$|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \leq \mathbb{P}(A)\phi(k).$$

Observe that we also have the trivial identity

$$\mathbb{P}\left(\overline{A} \cap \overline{B}\right) - \mathbb{P}\left(\overline{A}\right)\mathbb{P}\left(\overline{B}\right) = \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B).$$

So, withou a loss of generality, we can suppose $\mathbb{P}(A) \leq \frac{1}{2}$. Now, taking the supremum over A, B and p, we conclude the proof. □

2.3.4 ϕ-mixing property for Z-adic digits

As mentionned in the previous Subsection, there exists a good upper bound on the ϕ-mixing coefficients for the coordinates of $x \in \mathbb{X}$.

Proposition 2.3.5. *For $x = (x_j)_{j\geq 2} \in \mathbb{X}$ randomly chosen with law $\mathbb{P}$, we have for $k \geq 1$*

$$\phi(k) \leq \frac{2}{\varphi^{2k}}.$$

Our way to prove this result needs to introduce a parametrization of the Markov Chain. So, we define the function

$$\begin{array}{lllll} \psi : & \{0,1\} \times [0,1] & \longrightarrow \{0,1\} & \\ & (0,t) & \longmapsto 1 & \text{if } t < \frac{1}{\varphi^2} \\ & (0,t) & \longmapsto 0 & \text{if } t \geq \frac{1}{\varphi^2} \\ & (1,t) & \longmapsto 0. & \end{array}$$

Now, let $(U_j)_{j\geq 2}$ iid random variables following a uniform law on the real unit interval and we define the sequence $(y_j)_{j\geq 2}$ as

$$\left\{ \begin{array}{l} y_j = \psi(y_{j-1}, U_j) \quad \forall j \geq 2 \\ y_1 = 0 \end{array} \right. \tag{2.23}$$

Lemma 2.3.6. *The law of $(y_j)_{j\geq 2}$ is $\mathbb{P}$.*

Proof of Lemma 2.3.6. The transition matrix is the same as for $\mathbb{P}$ and so is the initial law : the probability that y_2 is 0 is the same as the event $U_2 \geq \frac{1}{\varphi^2}$ that is $\frac{1}{\varphi} = \mathbb{P}(C_0)$. □

Proof of Proposition 2.3.5. Let $k \geq 1$ and $p \geq 2$. To choose x with law $\mathbb{P}$ is equivalent to construct a sequence $y = (y_j)_{j\geq 2}$ using the process (2.23) thanks to Lemma 2.3.6. Thus, we have a sequence of iid random variables $(U_j)_{j\geq 2}$ with a uniform law on the real unit interval. We consider the event

$$C := \left\{ \exists j \in]p, p+k] : U_j \geq \frac{1}{\varphi^2} \right\}$$

which has a probability

$$\mathbb{P}(C) = 1 - \frac{1}{\varphi^{2k}} > 0.$$

If $y \in C$ then it implies that $y_j = 0$ because of (2.23). Then, we take $A \in \sigma(x_j : 2 \leq j \leq p)$ and $B \in \sigma(x_j : j \geq k + p)$. We observe that, due to (2.23), $A \in \sigma(U_j : 2 \leq j \leq p)$ while $C \in \sigma(U_j : p < j \leq k + p)$ thus A and C are independent. Also, due to Proposition 2.3.2, we observe $B \cap C \in \sigma(U_j : j > p)$. So A and $B \cap C$ are independent. Also, we have

$$\begin{aligned}|\mathbb{P}_A(B) - \mathbb{P}(B)| \leq |\mathbb{P}_A(B) - \mathbb{P}_{A\cap C}(B)| + |\mathbb{P}_{A\cap C}(B) - \mathbb{P}_C(B)| \\ + |\mathbb{P}_C(B) - \mathbb{P}(B)|.\end{aligned}$$

But the independences between A and C or A and $B \cap C$ give that

$$|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \leq |\mathbb{P}(A \cap B) - \mathbb{P}_C(A \cap B)| + |\mathbb{P}_C(B) - \mathbb{P}(B)|. \tag{2.24}$$

As shown in [6] (Lemma 4.3), denoting $\overline{C}$ the complement of C, we have the general inequality for any event D

$$|\mathbb{P}_C(D) - \mathbb{P}(D)| \leq \mathbb{P}(\overline{C}) \tag{2.25}$$

which implies in (2.24) that

$$|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \leq 2\mathbb{P}(\overline{C}).$$

□

Another property about the measure is that if indeed $\mathbb{P}(x_k = 1)$ depends on k, it is actually bounded between two positive values. We generalise that fact with the next proposition.

Proposition 2.3.7. *Let $(k_0, \cdots, k_\ell)$ be a collection of integers such that*

$$2 \leq k_0 < k_1 < \cdots < k_\ell$$

and also A be a union of cylinders of order k_0 such that $x_{k_0} = 0$ if $x \in A$ and such that $\mathbb{P}(A) > 0$. Then

$$\frac{1}{\varphi^\ell} \leq \mathbb{P}_A\ (\forall 1 \leq i \leq \ell\ :\ x_{k_i} = 0) \leq \left(\frac{2}{\varphi^2}\right)^\ell.$$

Proof. Let $I \subset \mathbb{N}$. Since there is a finite number of cylinders of order k_0, we can write A as a disjoint union $\sqcup_{i\in I} C^{(i)}$ where $C^{(i)}$ is a cylinder of order k_0. We have the identity

$$\mathbb{P}_A\ (\forall 1 \leq i \leq \ell\ :\ x_{k_i} = 0) = \prod_{i=1}^{\ell} \mathbb{P}\ (x_{k_i} = 0 \mid A \cap \bigcap_{j=1}^{i-1} (x_{k_j} = 0)) \tag{2.26}$$

with the convention $\cap_{j=1}^{0}(x_{k_j} = 0) = \mathbb{X}$. But we have

$$\begin{aligned}\mathbb{P}_A(x_{k_1} = 0) &= \frac{\mathbb{P}\left((x_{k_1} = 0) \cap \left(x \in \sqcup_{i\in I} C^{(i)}\right)\right)}{\mathbb{P}(A)} \\ &= \frac{1}{\mathbb{P}(A)} \sum_{i\in I} \mathbb{P}\left((x_{k_1} = 0) \cap x \in C^{(i)}\right).\end{aligned}$$

We recall $\mathbb{P}(A) = \frac{|I|}{\varphi^{k_0}}$ and that the summand is the probability of a cylinder of order k_1 with a digit 0 at the leftmost position so its measure is $\frac{1}{\varphi^{k_1}}$. Thus, we obtain

$$\mathbb{P}_A(x_{k_1} = 0) = \mathbb{P}(x_{k_1-k_0} = 0).$$

We proceed similarly for the other terms in (2.26) and get

$$\mathbb{P}_A\ (\forall 1 \leq i \leq \ell\ :\ x_{k_i} = 0) = \prod_{i=1}^{\ell} \mathbb{P}\ (x_{k_i-k_{i-1}} = 0). \tag{2.27}$$

The last equality is given using the renewal of $\mathbb{P}$. Now we claim that, for any $k \geq 2$

$$\frac{1}{\varphi} \leq \mathbb{P}(x_k = 0) \leq \frac{2}{\varphi^2}.$$

Indeed, we recall $\mathbb{P}(x_k = 0) = \frac{F_k}{\varphi^{k-1}}$ and observe that the subsequences $\left(\frac{F_{2k}}{\varphi^{2k-1}}\right)$ and $\left(\frac{F_{2k+1}}{\varphi^{2k}}\right)$ are adjacent sequences. □

2.3.5 Ergodic convergence

This subsection is exactly the same as the end of Subsection 2.1 in [6] in the context of an integer base but we write it again in the Z-adic context where the arguments are identical.

For x in $\mathbb{X}$, we define the sequence of empirical probability measures along the (beginning of the) orbit of x: for every $N \geq 1$, we set

$$\epsilon_N(x) := \frac{1}{N} \sum_{0 \leq n < N} \delta_{T^n x}$$

(where δ_y denotes the Dirac measure on $y \in \mathbb{X}$).
Since the space of probability measures on $\mathbb{X}$ is compact for the weak-$*$ topology, we can always extract a convergent subsequence. Moreover, every subsequential limit of $(\epsilon_N(x))$ is a T-invariant probability measure. By the uniqueness of the T-invariant probability measure, for every $x \in \mathbb{X}$ we have $\epsilon_N(x) \to \mathbb{P}$. In other words, we have the convergence

$$\forall x \in \mathbb{X},\ \ \forall f \in \mathcal{C}(\mathbb{X}), \qquad \frac{1}{N} \sum_{0 \leq n < N} f(T^n x) \xrightarrow[N \to +\infty]{} \int_{\mathbb{X}} f \mathrm{d}\mathbb{P}. \tag{2.28}$$

We will be interested here in the special case $x = 0$ because $\mathbb{N} = \{T^n 0 : n \in \mathbb{N}\}$. Then (2.28) becomes

$$\forall f \in \mathcal{C}(\mathbb{X}), \qquad \frac{1}{N} \sum_{0 \leq n < N} f(n) \xrightarrow[N \to +\infty]{} \int_{\mathbb{X}} f \mathrm{d}\mathbb{P}. \tag{2.29}$$

Equation (2.29) shows that, for a continuous function f, averaging f over $\mathbb{N}$ (for the natural density) amounts to averaging over $\mathbb{X}$ (for $\mathbb{P}$). The next section shows how this convergence can be extended to some non-continuous functions related to the sum-of-digits function.

2.4 Sum of digits on the Odometer

This section is very similar to the corresponding section in [6] for the integer base case. The only differences are the need of the new Lemma 2.4.4. All the other results in this section are the same even though we need to adapt the proofs. For every integer $k \geq 2$, we define the continuous map $s_k : \mathbb{X} \to \mathbb{Z}$ as the sum of the digits of indices $\leq k$, that is to say

$$s_k(x) := x_k + \cdots + x_2.$$

Let $r \in \mathbb{N}$. We define the functions $\Delta_k^{(r)} : \mathbb{X} \to \mathbb{Z}$ by

$$\Delta_k^{(r)}(x) := s_k(x + r) - s_k(x).$$

The functions $\Delta_k^{(r)}$ are well-defined, continuous (and bounded) on $\mathbb{X}$. By (2.29), we have

$$\frac{1}{N}\sum_{n<N} \Delta_k^{(r)}(n) = \frac{1}{N}\sum_{n<N} \Delta_k^{(r)}(T^n 0) \xrightarrow[N\to+\infty]{} \int_{\mathbb{X}} \Delta_k^{(r)} \mathrm{d}\mathbb{P}. \tag{2.30}$$

Although the sum-of-digits function s is not well defined on $\mathbb{X}$, we can extend the function $\Delta^{(r)}$ defined by (2.3) on the set of $x \in \mathbb{X}$ for which the number of different digits between x and $x + r$ is finite. This subset contains the Z-adic integers x such that there exists an index $k \geq 2 + \max(\{\ell : r_\ell \neq 0\})$ such that $x_k = x_{k+1} = 0$ (see Lemma 2.2.5). So, except for a finite number of Z-adic integers, we can define

$$\Delta^{(r)}(x) := \lim_{k\to\infty} \Delta_k^{(r)}(x).$$

Remark 6. *Let t, u be two integers. For every integer k we have the decomposition formula*

$$\Delta_k^{(t+u)} = \Delta_k^{(t)} + \Delta_k^{(u)} \circ T^t. \tag{2.31}$$

So, taking $\mathbb{P}$-almost everywhere the limit when k tends to infinity, we get

$$\Delta^{(t+u)} = \Delta^{(t)} + \Delta^{(u)} \circ T^t \ \ (\mathbb{P}\text{-}a\text{-}s.). \tag{2.32}$$

Then, by induction on t, we deduce

$$\Delta^{(t)} = \Delta^{(1)} + \Delta^{(1)} \circ T + \cdots + \Delta^{(1)} \circ T^{t-1} \ \ (\mathbb{P}\text{-}a\text{-}s.). \tag{2.33}$$

$\Delta^{(r)}$ is not bounded on $\mathbb{X}$ therefore it is not continuous. So, (2.29) is not applicable for $f \circ \Delta^{(r)}$ with f a continuous map on $\mathbb{X}$. However, it is possible to get the same convergence as in (2.29) with weaker assumptions on f than continuity.

Proposition 2.4.1. *Let $r \geq 1$ and $f : \mathbb{Z} \to \mathbb{C}$. Assume that there exist $\alpha \geq 1$ and C in $\mathbb{R}_+^*$ such that for every $n \in \mathbb{Z}$*

$$|f(n)| \leq C|n|^\alpha + |f(0)|. \tag{2.34}$$

Then $f \circ \Delta^{(r)} \in L^1(\mathbb{P})$ and we have the convergence

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N}\sum_{n<N} f(\Delta^{(r)}(n)) &= \int_{\mathbb{X}} f(\Delta^{(r)}(x))\mathrm{d}\mathbb{P}(x) \\ &= \lim_{k\to\infty} \int_{\mathbb{X}} f(\Delta_k^{(r)}(x))\mathrm{d}\mathbb{P}(x).\end{aligned}$$

Corollary 2.4.2. *For every $d \in \mathbb{Z}$*

$$\begin{aligned}\mu^{(r)}(d) &:= \lim_{N\to\infty} \frac{1}{N}\left|\left\{n < N : \Delta^{(r)}(n) = d\right\}\right| \\ &= \mathbb{P}\left(\left\{x \in \mathbb{X} : \Delta^{(r)}(x) = d\right\}\right). \end{aligned} \tag{2.35}$$

Moreover, $\Delta^{(r)}$ has zero-mean and has finite moments.

The proof of this corollary is exactly the same as in [6]. In particular, we just prove the existence of the asymptotic densities of the sets $\{n \in \mathbb{N} \mid \Delta^{(r)}(n) = d\}$, where $d \in \mathbb{Z}$. In integer base, it is easy to prove (see [3] or [6]). It is harder here to follow Besineau's proof because it uses the arithmetic properties of the sum-of-digits (in integer base) function that we do not have anymore.

Remark 7. *Using trivial arguments, Proposition 2.4.1 and Corollary 2.4.2 are also true when $r = 0$. We observe that $\mu^{(0)} = \delta_0$.*

Before proving this proposition and its corollary, we need the following lemma that looks like Lemma 1.29 in [26].

Lemma 2.4.3. *Let $r \geq 1$. For $N \geq 1$, $k \geq 2$ and $d, d' \in \mathbb{Z}$, we have the inequality*

$$\begin{aligned}&\frac{1}{N}\left|\{n < N : (\Delta^{(r)}(n), \Delta_k^{(r)}(n)) = (d, d')\}\right| \\ &\qquad\qquad \leq r\varphi^3\ \mathbb{P}\Big(\{x \in \mathbb{X} : (\Delta^{(r)}(x), \Delta_k^{(r)}(x)) = (d, d')\}\Big).\end{aligned} \tag{2.36}$$

In particular, we have

$$\frac{1}{N}\left|\{n < N : \Delta^{(r)}(n) = d\}\right| \leq r\varphi^3\ \mathbb{P}\Big(\{x \in \mathbb{X} : \Delta^{(r)}(x) = d\}\Big). \tag{2.37}$$

Proof of Lemma 2.4.3. We adapt the proof of Lemma 2.3 in [6]. First, (2.36) implies (2.37) so we just prove (2.36). We fix $k \geq 2$. For $\ell \geq 1$, let V_ℓ be the set of the values reached by the couple $(\Delta^{(r)}, \Delta_k^{(r)})$ on the $F_{\ell+1} - r$ (resp. $F_\ell - r$) first levels of the large (resp. small) tower of order ℓ. Of course, if $F_{\ell+1} - r \leq 0$ then $V_\ell := \emptyset$. In particular, for any $r \geq 1$, $V_1 = \emptyset$. Also, if $F_\ell \leq r < F_{\ell+1}$ then V_ℓ is defined considering only the large tower of order ℓ. For any ℓ, we observe that V_ℓ is a finite set. On each level that is not in the r top levels of the large or small tower, the first ℓ digits of both x and $x + r$ are constant, and digits of higher order are the same. Therefore, $\Delta^{(r)}$ and $\Delta_k^{(r)}$ are constant on such a level. We observe that the sequence $(V_\ell)_{\ell\geq 1}$ is increasing for the inclusion.

Now, for $d, d' \in \mathbb{Z}$, there are 2 cases.

1. If $(d, d') \notin \cup_{\ell\geq 2} V_\ell$, then for each $n \in \mathbb{N}$ we have $(\Delta^{(r)}(n), \Delta_k^{(r)}(n)) \neq (d, d')$. Indeed, for each $n \in \mathbb{N}$, there exists a smallest integer $\ell \geq 2$ such that n is in the first $F_{\ell+1} - r$ levels of the large tower of order ℓ, hence $(\Delta^{(r)}(n), \Delta_k^{(r)}(n)) \in V_\ell$. In this case, (2.36) is trivial.

2. If $(d, d') \in \cup_{\ell\geq 1} V_\ell$ then there exists a unique $\ell \geq 2$ such that $(d, d') \in V_\ell \backslash V_{\ell-1}$. We observe that, due to the Cut and Stack process, the value (d, d') must appear firstly in the large tower of order ℓ. Since $(\Delta^{(r)}, \Delta_k^{(r)})$ is constant on each of the first $F_{\ell+1} - r$ levels of the large tower, it takes the value (d, d') on at least one whole such level which is of measure $\frac{1}{\varphi^\ell}$. So, we have

$$\mathbb{P}\Big(\{x \in \mathbb{X} : (\Delta^{(r)}(x), \Delta_k^{(r)}(x)) = (d, d')\}\Big) \geq \frac{1}{\varphi^\ell}. \tag{2.38}$$

Also, since the couple $(d, d') \notin V_{\ell-1}$, we claim that, for every $N \geq 1$

$$\frac{1}{N}\left|\{n < N : (\Delta^{(r)}(n), \Delta_k^{(r)}(n)) = (d, d')\}\right| \leq \frac{r}{F_{\ell-1}}.$$

Indeed,

(a) If $r \geq F_{\ell-1}$, the inequality is then trivial.

(b) If $r < F_{\ell-1}$ then, since (d, d') is not in $V_{\ell-1}$, (d, d') can only appear inside a part the r highest levels of the big or the small towers of order $\ell - 1$. Let us denote C the union of these r highest levels. Since 0 lies in the bottom level of the large tower of order $\ell - 1$, the set S of integers $n \geq 0$ such that $T^n 0 \in C$ has the following properties :

- $\{0, \cdots, F_\ell - r - 1\} \cap S = \emptyset$ and
- S is the union of subsets formed by r consecutive integers separated by gaps of length $F_\ell - r$ or $F_{\ell-1} - r$.

So, we have

$$\begin{aligned}\frac{1}{N}\left|\{0 \leq n < N : (\Delta^{(r)}(n), \Delta_k^{(r)}(n)) = (d, d')\}\right| & \\ \leq \frac{1}{N}\left|\{0 \leq n < N : T^n 0 \in C\}\right| & \\ \leq \frac{r}{F_{\ell-1}}. &\end{aligned}$$

Since $F_{\ell-1} \geq \varphi^{\ell-3}$ (by double induction), we get

$$\frac{1}{N}\left|\{n < N : (\Delta^{(r)}(n), \Delta_k^{(r)}(n)) = (d, d')\}\right| \leq \frac{r}{\varphi^{\ell-3}}. \tag{2.39}$$

Combining inequalities (2.38) and (2.39) gives (2.36).

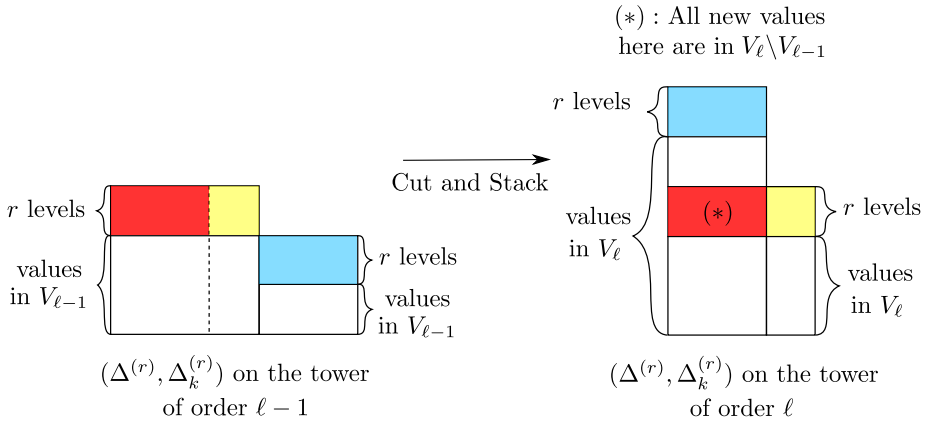


Figure 2.6: Visual description of $V_{\ell-1}$, V_ℓ and $V_\ell \backslash V_{\ell-1}$.

□

As shown in (2.33), the understanding of $\Delta^{(1)}$ (and so $\Delta_k^{(1)}$ for $k \geq 2$) is fundamental to understand $\Delta^{(r)}$ so we also state the following lemma.

Lemma 2.4.4. *The function $\Delta^{(1)}$ is well-defined for $x \in \mathbb{X}$ if and only if $x \in C_{00(10)^d} \cup C_{001(01)^d}$ for some $d \geq 0$.*

Furthermore, if $x \in C_{00(10)^d}$ then $\Delta^{(1)}(x) = 1 - d$ and

$$\Delta_k^{(1)}(x) = \begin{cases} \frac{2-k}{2} & \text{if } k \equiv 0[2] \text{ and } k \leq 2d, \\ \frac{1-k}{2} & \text{if } k \equiv 1[2] \text{ and } k \leq 2d+1, \\ 1-d & \text{if } k \geq 2d+2. \end{cases}$$

Also, if $x \in C_{001(01)^d}$ then $\Delta^{(1)}(x) = -d$ and

$$\Delta_k^{(1)}(x) = \begin{cases} \frac{-k}{2} & \text{if } k \equiv 0[2] \text{ and } k \leq 2d+2, \\ \frac{1-k}{2} & \text{if } k \equiv 1[2] \text{ and } k \leq 2d+1, \\ \frac{3-k}{2} & \text{if } k = 2d+3, \\ -d & \text{if } k \geq 2d+4. \end{cases}$$

Proof of Lemma 2.4.4. If $x \in C_{00(10)^d}$ for some $d \geq 0$, the definition of T gives that $x+1 \in C_{01(00)^d}$ with the left digits unchanged. Thus the sequence $(\Delta_k^{(1)}(x))$ is stationary, $\Delta^{(1)}(x)$ is well defined and $\Delta^{(1)}(x) = 1 - d$. Also, if $x \in C_{001(01)^d}$ for some $d \geq 0$, then $x + 1 \in C_{010(00)^d}$, $\Delta^{(1)}(x)$ is well defined and $\Delta^{(1)}(x) = -d$. If x does not belong to those cylinders then $x = (01)^\infty$ or $x = (10)^\infty$. In that case, $x + 1 = \overline{0^\infty}$ and the sequence $(\Delta_k^{(1)}(x))$ diverges to $-\infty$. This proves the equivalence. The values of $\Delta_k^{(1)}(x)$ can be easily found by looking at the addition we write in 2.2.1. □

Proof of Proposition 2.4.1. We adapt the proof of Proposition 2.1 in [6]. Let $\varepsilon > 0$. For any integer $k \geq 2$, we have

$$\left| \frac{1}{N} \sum_{n<N} f(\Delta^{(r)}(n)) - \int_{\mathbb{X}} f(\Delta^{(r)}(x)) \mathrm{d}\mathbb{P}(x) \right| \leq A_1 + A_2 + A_3,$$

where

$$A_1 := \frac{1}{N} \sum_{n<N} \left| f(\Delta^{(r)}(n)) - f(\Delta_k^{(r)}(n)) \right|,$$

$$A_2 := \left| \frac{1}{N} \sum_{n<N} f(\Delta_k^{(r)}(n)) - \int_{\mathbb{X}} f(\Delta_k^{(r)}(x)) \mathrm{d}\mathbb{P}(x) \right|,$$

$$A_3 := \int_{\mathbb{X}} \left| f(\Delta^{(r)}(x)) - f(\Delta_k^{(r)}(x)) \right| \mathrm{d}\mathbb{P}(x).$$

However, as in [6], we have

$$\begin{aligned}
A_1 &= \frac{1}{N}\sum_{n<N}\sum_{j,j'\in\mathbb{Z}}\Big|f(j)-f(j')\Big|\mathbb{1}_{(j,j')}\Big(\Delta^{(r)}(n),\Delta_k^{(r)}(n)\Big)\\
&= \sum_{j,j'\in\mathbb{Z}}\Big|f(j)-f(j')\Big|\frac{1}{N}\sum_{n<N}\mathbb{1}_{(j,j')}\Big(\Delta^{(r)}(n),\Delta_k^{(r)}(n)\Big)\\
&\leq rb\sum_{j,j'\in\mathbb{Z}}\Big|f(j)-f(j')\Big|\,\mathbb{P}\Big(\Delta^{(r)}(n)=j,\Delta_k^{(r)}(n)=j'\Big)\\
&= rb\int_{\mathbb{X}}\Big|f(\Delta^{(r)}(x))-f(\Delta_k^{(r)}(x))\Big|\mathrm{d}\mathbb{P}(x)=rbA_3.
\end{aligned}$$

It follows that

$$\left|\frac{1}{N}\sum_{n<N}f(\Delta^{(r)}(n))-\int_{\mathbb{X}}f(\Delta^{(r)}(x))\mathrm{d}\mathbb{P}(x)\right|\leq A_2+(1+rb)A_3.$$

We want to apply a Dominated Convergence Theorem to deal with A_3 (observe that we have $f(\Delta_k^{(r)}(x))\xrightarrow[k\to\infty]{}f(\Delta^{(r)}(x))$ $\mathbb{P}$-almost-surely). For this, we need to find a good dominant function. As in [6], we define $g_i:=\sup_{k\geq 2}|\Delta_k^{(1)}\circ T^i(x)|$ for $i=0,\cdots,r-1$ and by (2.34) we get the inequalities

$$\Big|f\circ\Delta_k^{(r)}(x)\Big|\leq C\sum_{j_0+\cdots+j_{r-1}=\alpha}\sum_{i=0}^{r-1}\frac{\binom{\alpha}{j_0,\cdots,j_{r-1}}}{r}g_i(x)^{rj_i}+\Big|f(0)\Big|$$

and

$$\Big|f\circ\Delta^{(r)}(x)\Big|\leq C\sum_{j_0+\cdots+j_{r-1}=\alpha}\sum_{i=0}^{r-1}\frac{\binom{\alpha}{j_0,\cdots,j_{r-1}}}{r}g_i(x)^{rj_i}+\Big|f(0)\Big|.$$

We need to prove that $g_i^{rj_i}$ is integrable for the measure $\mathbb{P}$. It is equivalent to show that $\sum_m\mathbb{P}(\{x\in\mathbb{X}:g_i^{rj_i}(x)>m\})$ is a convergent series. We have

$$\begin{aligned}
g_i(x)>m^{\frac{1}{rj_i}} &\Leftrightarrow \sup_{k\in\mathbb{N}}|\Delta_k^{(1)}\circ T^i(x)|>m^{\frac{1}{rj_i}}\\
&\Leftrightarrow \exists k\in\mathbb{N},\ \ |\Delta_k^{(1)}(T^ix)|>m^{\frac{1}{rj_i}}.
\end{aligned}$$

From Lemma 2.4.4

$$\exists k\in\mathbb{N},\ \ |\Delta_k^{(1)}(T^ix)|>m^{\frac{1}{rj_i}}\Leftrightarrow T^ix\in C_{00(10)^{\lfloor m^{\frac{1}{rj_i}}\rfloor}}\cup C_{0(01)^{\lfloor m^{\frac{1}{rj_i}}\rfloor}}$$

It follows, by T-invariance

$$\mathbb{P}(\{x\in\mathbb{X}:g_i^{rj_i}(x)>m\})\leq\left(\frac{1}{\varphi}\right)^{2+2\lfloor m^{\frac{1}{rj_i}}\rfloor}+\left(\frac{1}{\varphi}\right)^{1+2\lfloor m^{\frac{1}{rj_i}}\rfloor}=\left(\frac{1}{\varphi}\right)^{2\lfloor m^{\frac{1}{rj_i}}\rfloor}.$$

The quantity on the RHS is the general term of a convergent series which shows that $g_i^{rj_i}$ is integrable for the measure $\mathbb{P}$. The dominated theorem can be applied and, for k large enough, $(1+rb)A_3 \leq \frac{\varepsilon}{2}$ for every $N \geq 1$. Now, once we have fixed such a k, for N large enough, A_2 is bounded by $\frac{\varepsilon}{2}$ because of (2.29) and the continuity of $\Delta_k^{(r)}$ and f. The convergence in the statement is thus proved.

Note that the argument of the dominated convergence theorem also proves that $f \circ \Delta^{(r)} \in L^1(\mathbb{P})$ and $\int_{\mathbb{X}} f \circ \Delta^{(r)} \mathrm{d}\mathbb{P} = \lim_{k\to\infty} \int_{\mathbb{X}} f \circ \Delta_k^{(r)} \mathrm{d}\mathbb{P}$. □

More generally, we have the following convergence.

Proposition 2.4.5. *Let $r \geq 1$ and $f : \mathbb{Z} \to \mathbb{C}$ be such that $f \circ \Delta^{(r)} \in L^1(\mathbb{P})$. Then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n<N} f(\Delta^{(r)}(n)) = \int_{\mathbb{X}} f(\Delta^{(r)}(x)) \mathrm{d}\mathbb{P}(x).$$

The proof is exactly the same as Proposition 2.4 in [6]. We have shown that the random variable $\Delta^{(r)}$ satisfies some good properties such as the finiteness of moments of any order of $\mu^{(r)}$, the law of $\Delta^{(r)}$. The next section focuses on another natural question : how to compute the law $\mu^{(r)}$?

2.5 How to compute $\mu^{(r)}$

In this section, we present an algorithm that computes exactly the measure $\mu^{(r)}$ for any r in $\mathbb{N}$ and the consequences of it.

2.5.1 Description of the algorithm

This algorithm is different from the classical way to compute $\mu^{(r)}$ in integer base (see [3], page 14 or [6] Proposition 3.1 for instance). In integer base $b \geq 2$, the computation of $\mu^{(r)}$ relies on inductive relations on the expansion of r which are easy to prove in that case. For instance, one of the relation states that if r is a multiple of b then the variation when adding r to x is the same as the variation when adding $\frac{r}{b}$ to σx. So the laws of the variation when adding r or $\frac{r}{b}$ are the same. In the Zeckendorf representation, this trivial argument is false because the carries can modify the digits on both sides (unlike the integer base case where only the left side can be affected). However, we find an algorithm to compute $\mu^{(r)}$, in this Zeckendorf system, that relies on Rokhlin towers. We precise that this algorithm can be adapted in integer base.

First, let $r \geq 1$ ($r = 0$ is irrelevant) and let $\ell \geq 2$ be the unique integer such that $F_\ell \leq r < F_{\ell+1}$. At a given order $k \geq 1$, we define CST_k as the union of the levels of the large and small towers of order k, except the r highest levels. We observe $\mathrm{CST}_1 = \emptyset$. On each level of CST_k, $\Delta^{(r)}$ is constant (see Proof of Lemma 2.4.3). But, due to the Cut-and-Stack process, the value of $\Delta^{(r)}$ on these "constant" levels of CST_k may be deduced from those taken in CST_{k-1}. That is why we also introduce, for $k \geq 2$, the *New information zone of the Rokhlin tower of order k* as the set

$$\mathrm{NIZ}_k := \mathrm{CST}_k \backslash \mathrm{CST}_{k-1}.$$

Of course, everything depends on r but we do not emphasise on that in the notations of ℓ, NIZ_k and CST_k for simplicity. Let us define clearly this set NIZ_k.

Proposition 2.5.1. *In a large tower, we enumerate, starting by 1, the levels from the base of the tower to the top of it. We have the following description of NIZ_k.*

- *If* $2 \leq k \leq \ell - 1$, $NIZ_k = \emptyset$.
- NIZ_ℓ *is the union of the* $F_{\ell+1} - r$ *first levels of the large tower of order* ℓ.
- $NIZ_{\ell+1}$ *is the union of the* F_ℓ *levels between the* $F_{\ell+1} - r + 1^{th}$ *and the* $F_{\ell+2} - r^{th}$ *levels of the large tower of order* $\ell + 1$.
- *If* $k > \ell + 1$, NIZ_k *is the union of the* r *levels between the* $F_k - r + 1^{th}$ *and the* F_k^{th} *levels of the large tower of order* k.

Remark 8. *By definition, for every* $k \geq 1$, *the set* NIZ_k *is either empty or a disjoint union of cylinders of order* k *(exactly* $F_{\ell+1} - r$ *if* $k = \ell$, F_ℓ *if* $k = \ell + 1$ *or* r *if* $k \geq \ell + 2$*). Moreover, we observe that* $(NIZ_k)_{k \geq \ell}$ *is a partition of* $\mathbb{X} \backslash \{(01)^\infty \cup (10)^\infty\}$.

Example 2.5.2. *Here, an example with* $r = 4 = \overline{101}$.

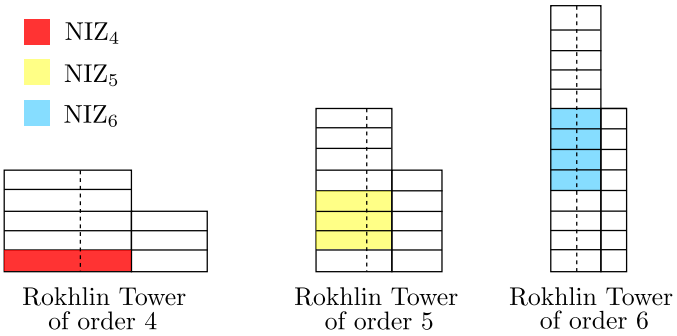


Figure 2.7: Visualization of $(\text{NIZ}_k)_{k \geq 1}$ when $r = 4 = \overline{101}$.

We also provide a more general figure to understand.

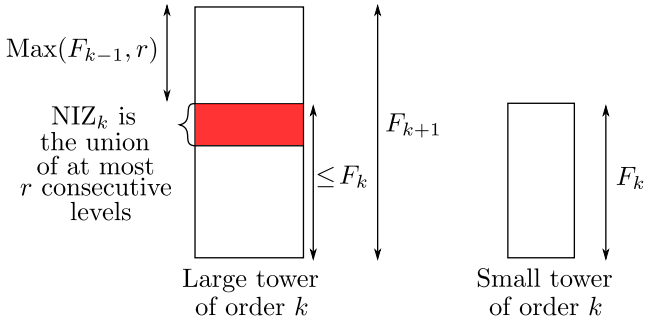


Figure 2.8: Visualisation of NIZ_k, $k \geq \ell + 2$.

Through the following lemmas, we state some observations on this sequence $(\text{NIZ}_k)_{k \geq 1}$. We start by a description of the cylinders of order k that appear in NIZ_k for $k \geq \ell + 2$.

Lemma 2.5.3. *If* $k \geq 2$, *the set* NIZ_k *is composed of cylinders of order* k *whose name has a double* 00 *at the leftmost position.*

Proof. Let $k \geq \ell + 2$, consider a cylinder $C_{n_{k+1}\cdots n_2}$ of order k that belongs in NIZ_k and where $(n_{k+1}, \cdots, n_2) \in \mathbb{X}_f$. Since this cylinder is part of the Rokhlin tower of order k, $n_{k+1} = 0$. Moreover, this cylinder is not in the F_{k-1} highest cylinder so $n_k = 0$ too. □

Now, we show a relation between cylinders of order k in NIZ_k and those of order $k+2$ in NIZ_{k+2} for $k \geq \ell$.

Lemma 2.5.4. *Let $k \geq 2$ and let $(n_{k-1}, \cdots, n_2) \in \mathbb{X}_f$. We have the implication*

$$C_{00n_{k-1}\cdots n_2} \subset NIZ_k \implies C_{0010n_{k-1}\cdots n_2} \subset NIZ_{k+2} \tag{2.40}$$

Remark 9. *If $k = 2$, we agree that the word $n_{k-1}\cdots n_2$ is the empty word.*

Proof. For $r = 1$, we let the reader check that $C_{00} \subset \mathrm{NIZ}_2$ and $C_{0010} \subset \mathrm{NIZ}_4$. For $r \geq 2$, we have $\mathrm{NIZ}_2 = \emptyset$. So, we can assume $k \geq 3$. Let $C_{00n_{k-1}\cdots n_2} \subset \mathrm{NIZ}_k$ and $n := \sum_{i=2}^{k-1} n_i F_i$. Then, $n \in \mathrm{NIZ}_k \cap [0, F_k[$ where the interval is an integer interval and considered as being part of $\mathbb{X}$.

1. If $k \geq \ell + 2$, we observe that $\mathrm{NIZ}_k \cap [0, F_k[= [F_k - r, F_k - 1[$. It follows that $F_{k+1} + n$, whose Zeckendorf expansion is $\overline{10n_{k-1}\cdots n_2}$, belongs to $[F_{k+2} - r, F_{k+2} - 1] \subset \mathrm{NIZ}_{k+2}$. Furthermore $F_{k+1} + n \in C_{0010n_{k-1}\cdots n_2}$ which is a level of the large tower of order $k+2$. So $C_{0010n_{k-1}\cdots n_2} \subset \mathrm{NIZ}_{k+2}$.

2. For $k = \ell$, we observe that $\mathrm{NIZ}_\ell \cap [0, F_\ell[= [0, F_{\ell+1} - r - 1[$. It follows that $n + F_{\ell+1} \in [F_{\ell+1}, 2F_{\ell+1} - r - 1]$. We claim that

$$F_{\ell+2} - r \leq F_{\ell+1} \leq 2F_{\ell+1} - r - 1 \leq F_{\ell+2} - 1. \tag{2.41}$$

   Indeed, since $F_\ell \leq r < F_{\ell+1}$;

   - $F_{\ell+2} - r \leq F_{\ell+2} - F_\ell = F_{\ell+1}$,
   - $F_{\ell+1} - r - 1 \geq 0$ and we deduce $2F_{\ell+1} - r - 1 \geq F_{\ell+1}$,
   - $2F_{\ell+1} - r - 1 \leq F_{\ell+1} + F_{\ell-1} - 1 \leq F_{\ell+2} - 1$.

   From (2.41), we get $n + F_{\ell+1} \in [F_{\ell+2} - r, F_{\ell+2} - 1]$. Since $[F_{\ell+2} - r, F_{\ell+2} - 1] \subset \mathrm{NIZ}_{\ell+2}$, we conclude as in the first point.

3. If $k = \ell + 1$, we get $\mathrm{NIZ}_k \cap [0, F_\ell[= [F_{\ell+1} - r, F_{\ell+2} - r - 1]$. It follows that $n + F_{\ell+2} \in [F_{\ell+3} - r, 2F_{\ell+2} - r - 1]$. Since $r \geq F_\ell$, $2F_{\ell+2} - r - 1 \leq F_{\ell+3} - 1$ and so $n + F_{\ell+2} \in [F_{\ell+3} - r, F_{\ell+3} - 1] \subset \mathrm{NIZ}_{\ell+2}$. We conclude as in the first point.

□

Lemma 2.5.5. *Let $k \geq \ell$ and a collection $(n_{k-1}, \cdots, n_2) \in \mathbb{X}_f$ such that $C_{00n_{k-1}\cdots n_2}$ is a cylinder of NIZ_k. Let $d \in \mathbb{Z}$ such that*

$$\Delta^{(r)}_{|C_{00n_{k-1}\cdots n_2}} = d$$

then

$$\Delta^{(r)}_{|C_{0010n_{k-1}\cdots n_2}} = d - 1.$$

These lemmas imply the following corollary that will give birth to our algorithm.

Corollary 2.5.6. *Let $k \geq \ell + 4$, the values taken by $\Delta^{(r)}$ on the cylinders of order $k+2$ that compose NIZ_{k+2} are exactly the values taken on cylinders of order k in NIZ_k shifted by -1.*

Proof of Lemma 2.5.5. Let $n := \sum_{i=2}^{k-1} n_i F_i$ such that $C_{00n_{k-1}\cdots n_2} \subset \mathrm{NIZ}_k$. We have $C_{00n_{k-1}\cdots n_2} \subset C_{0n_{k-1}\cdots n_2}$ which is, due to the Cut and Stack process, one of the r top levels of the large tower of order $k-1$. Also, due to the Cut and Stack process, $n+r$ must lies in the small tower of order $k-1$ (see Figure 2.9) so $(n+r)_k$ must be 1.

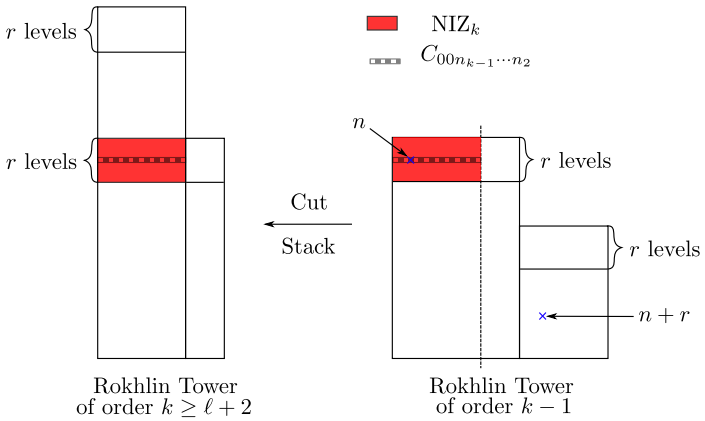


Figure 2.9: Main argument to justify $(n+r)_k = 1$.

So, we have the following addition

(n)		0	0	n_{k-1}	$\cdots$	n_ℓ	$\cdots$	n_2
(r)	$+$					r_ℓ	$\cdots$	r_2
$(n+r)$	$=$	0	1	0	$\cdots$	$(n+r)_\ell$	$\cdots$	$(n+r)_2$

with, by hypothesis, $\Delta^{(r)}(n) = d$. Then $n + F_{k+1} + r$ is calculated as follows

$(n+F_{k+1})$		0	0	1	0	n_{k-1}	$\cdots$	n_ℓ	$\cdots$	n_2
(r)	$+$							r_ℓ	$\cdots$	r_2
$(n+F_{k+1}+r)$	$=$	0	0	1	1	0	$\cdots$	$(n+r)_\ell$	$\cdots$	$(n+r)_2$
$(n+F_{k+1}+r)$	$=$	0	1	0	0	0	$\cdots$	$(n+r)_\ell$	$\cdots$	$(n+r)_2$

The sum of digits of $n + F_{k+1} + r$ is the same as those of $n + r$ shifted by -1 due to the correction of the expansion made. Hence, we get $\Delta^{(r)}(n + F_{k+1}) = \Delta^{(r)}(n) - 1$ i.e.

$$\Delta^{(r)}_{|C_{0010n_{k-1}\cdots n_2}} = \Delta^{(r)}_{|C_{00n_{k-1}\cdots n_2}} - 1.$$

□

Proof of Corollary 2.5.6. Definition 2.5.1 ensures that the same number of cylinders of order $k+2$ in NIZ_{k+2} is equal to the number of cylinders of order k in NIZ_k : there are r cylinders of each order. Lemma 2.5.4 and Lemma 2.5.5 enable us to conclude. □

These results enable us to give an algorithm that will compute the value of $\mu^{(r)}(d)$ for any $d \in \mathbb{Z}$. This algorithm has a low complexity (polynomial). Indeed, Corollary 2.5.6 implies that we can anticipate the values that will appear in the New Information Zone of order k. So we just need to compute $\Delta^{(r)}$ on a finite number of integers to compute exactly $\mu^{(r)}(d)$ ($d \in \mathbb{Z}$). We will precise this fact later in this Section. Before, we give a pseudo-code and provide an example of computation.

2.5.2 Pseudo-code

Algorithm 1 Compute $\mu^{(r)}(d)$

Require: $r \geq 1$ and $d \in \mathbb{Z}$
Ensure: $\mu^{(r)}(d)$
  STEP 0: Find the unique $\ell \geq 1$ such that $F_\ell \leq r < F_{\ell+1}$
  STEP 1: Passage in NIZ_ℓ
  Define $App_\ell \leftarrow 0$ (number of apparition of d during the for-loop)
  for $n = 0, \cdots, F_{\ell+1} - r - 1$ **do**
    if $\Delta^{(r)}(n) = d$ **then**
      $App_\ell \leftarrow App_\ell + 1$
    end if
  end for
  STEP 2: Passage in $\text{NIZ}_{\ell+1}$
  Define $App_{\ell+1} \leftarrow 0$
  for $n = F_{\ell+1} - r \cdots, F_{\ell+2} - r - 1$ **do**
    if $\Delta^{(r)}(n) = d$ **then**
      $App_{\ell+1} \leftarrow App_{\ell+1} + 1$
    end if
  end for
  STEP 3: Passage in $\text{NIZ}_{\ell+2}$
  Define $App_{\ell+2} \leftarrow 0$
  for $n = F_{\ell+2} - r, \cdots, F_{\ell+2} - 1$ **do**
    Store $\Delta^{(r)}(n)$ in an array **ARR1** (with multiplicity)
    if $\Delta^{(r)}(n) = d$ **then**
      $App_{\ell+2} \leftarrow App_{\ell+2} + 1$
    end if
  end for
  STEP 4: Passage in $\text{NIZ}_{\ell+3}$
  Define $App_{\ell+3} \leftarrow 0$
  for $n = F_{\ell+3} - r, \cdots, F_{\ell+3} - 1$ **do**
    Store $\Delta^{(r)}(n)$ in an array **ARR2** (with multiplicity)
    if $\Delta^{(r)}(n) = d$ **then**
      $App_{\ell+3} \leftarrow App_{\ell+3} + 1$
    end if
  end for

STEP 5: Search for d in higher orders
for $i \in$ **ARR1 do**
  if $i > d$ **then**
    d will appear at the order $\ell + 2 + 2(i-d)$ so
    if $App_{\ell+2+2(i-d)}$ is defined **then**
      $App_{\ell+2+2(i-d)} \leftarrow App_{\ell+2+2(i-d)} + 1$
    else
      $App_{\ell+2+2(i-d)} \leftarrow 1$
    end if
  end if
end for
for $i \in$ **ARR2 do**
  if $i > d$ **then**
    d will appear at the order $\ell + 3 + 2(i-d)$ so
    if $App_{\ell+3+2(i-d)}$ is defined **then**
      $App_{\ell+3+2(i-d)} \leftarrow App_{\ell+3+2(i-d)} + 1$
    else
      $App_{\ell+3+2(i-d)} \leftarrow 1$
    end if
  end if
end for
STEP 6: Compute $\mu^{(r)}(d)$
$\mu^{(r)}(d) \leftarrow \sum_{k \geq \ell} \frac{App_k}{\varphi^k}$

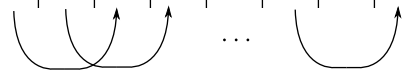
2.5.3 An example : computation of $\mu^{(4)}$

Let us compute the measure $\mu^{(r)}$ for $r = 4$. We have $\ell = 4$. Looking at the Rokhlin towers (see Example 2.5.2) gives that

- $\mathrm{NIZ}_k = \emptyset$ if $k = 1, 2, 3$,
- $\mathrm{NIZ}_4 = C_{0000}$,
- $\mathrm{NIZ}_5 = C_{00001} \sqcup C_{00010} \sqcup C_{00100}$,
- $\mathrm{NIZ}_6 = C_{000101} \sqcup C_{001000} \sqcup C_{001001} \sqcup C_{001010}$,
- $\mathrm{NIZ}_7 = C_{0010000} \sqcup C_{0010001} \sqcup C_{0010010} \sqcup C_{0010100}$.

The next terms of NIZ_k are found using Lemma 2.5.4. Using Corollary 2.5.6, one can construct the following table that gives the values taken by $\Delta^{(r)}$ (with multiplicity) on each cylinder of NIZ_k.

Order k	4	5	6	7	8	9	10	11	12	13
Values of			0	−1	−1	−2	−2	−3	−3	−4
$\Delta^{(4)}$ on		1	0	0	−1	−1	−2	−2	−3	−3
the levels	2	1	1	0	0	−1	−1	−2	−2	−3
of NIZ_k		0	−1	−1	−2	−2	−3	−3	−4	−4



Values in the columns on both sides of the arrows are shifted by -1

Figure 2.10: Table of the values of $\Delta^{(4)}$ on NIZ_k.

Thus, it is possible to know how many times some integer $d \in \mathbb{Z}$ appear in each column (ie at each order) and, since we know the measure of a cylinder at each order, it is possible to deduce the value of $\mu^{(r)}(d)$. For instance

- The value 2 appears once at order 4 so $\mu^{(4)}(2) = \frac{1}{\varphi^4}$.
- The value 1 appears twice at order 5, once at order 6 so $\mu^{(4)}(1) = \frac{2}{\varphi^5} + \frac{1}{\varphi^6}$.
- The value 0 appears once at order 5 and 8. It also appears twice at order 6 and 7, we deduce that $\mu^{(4)}(0) = \frac{1}{\varphi^5} + \frac{2}{\varphi^6} + \frac{2}{\varphi^7} + \frac{1}{\varphi^8}$.
- The value -1 appears once at order 6 and 10. It appears twice at order 7, 8 and 9. We deduce $\mu^{(4)}(-1) = \frac{1}{\varphi^6} + \frac{2}{\varphi^7} + \frac{2}{\varphi^8} + \frac{2}{\varphi^9} + \frac{1}{\varphi^{10}}$.
- Since -1 is a value at order $\ell + 2 = 6$ that does not appear in the previous order, we deduce by Corollary 2.5.6 that the value -2 will appear as many times as -1 appear but at orders incremented by 2. In other words, the value -2 appears once at order 8 and 12. It appears twice at order 9, 10 and 11. Thus $\mu^{(4)}(-2) = \mu^{(4)}(-1) \times \frac{1}{\varphi^2}$. We observe that -2 does not appear in the lower orders. This observation will be better explained in Proposition 2.5.7 and Corollary 2.1.1.

So, we get after simplification

$$\mu^{(4)}(d) = \begin{cases} 0 & \text{if } d > 2 \\ \frac{1}{\varphi^4} & \text{if } d = 2 \\ \frac{1}{\varphi^3} & \text{if } d = 1 \\ \frac{2}{\varphi^4} & \text{if } d = 0 \\ \mu^{(4)}(-1) \times \varphi^{2d+2} & \text{if } d < 0 \end{cases}$$

(where $\mu^{(4)}(-1) = \frac{1}{\varphi^4} + \frac{1}{\varphi^6}$.)

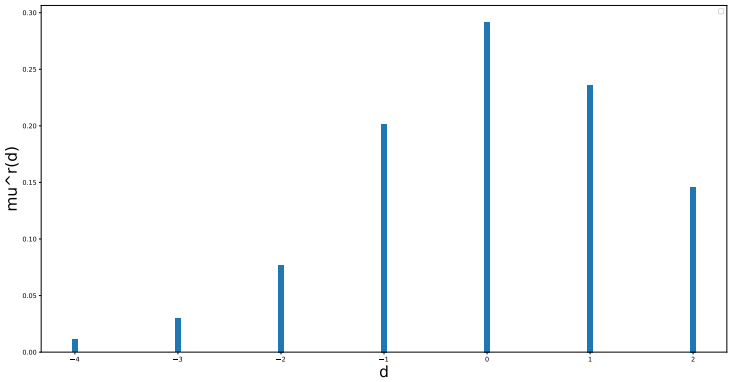


Figure 2.11: Bar chart of $\mu^{(4)}$.

As an exercise, the reader can check that $\mu^{(1)} = \mu^{(F_3)} = \mu^{(F_4)}$. In the next Section, we are going to prove that, for any $k \geq 2$

$$\mu^{(F_k)} = \mu^{(1)}.$$

2.5.4 Remarks on the algorithm

Once again, the algorithm described in Subsections 2.5.1 and 2.5.2 can be adapted in integer base so the following observations can also be adapted in integer base. Also, we write an algorithm that returns, for a given $r \geq 1$ and $d \in \mathbb{Z}$, the value $\mu^{(r)}(d)$. Of course, the algorithm can also be adapted to return $\mu^{(r)}(d)$ for every $d \in \mathbb{Z}$ or, at least to be executed on a computer, for a finite number of integers d without more computations (the adaptation only consist in manipulations of arrays).

In this subsection, let $r \geq 1$. Let also $\ell \geq 1$ be the unique integer such that $F_\ell \leq r < F_{\ell+1}$. The first important observation due to the algorithm is that we only need to compute $\Delta^{(r)}(n)$ for a finite number of n's to know exactly the quantity $\mu^{(r)}$. More precisely, we need to compute $\Delta^{(r)}$ for

- $F_{\ell+1} - r$ integers during **STEP 1**,
- F_ℓ integers during **STEP 2**,
- r integers during **STEP 3** and
- r integers during **STEP 4**.

So, we need to compute the image of $F_{\ell+2} + r$ different integers to compute $\mu^{(r)}$ exactly. One can prove that, when $r = F_\ell$, it is possible to adapt the algorithm to reduce that number to $F_{\ell+2}$. The adaptation consists into forgetting **STEP 4** and store the F_ℓ values in **STEP 2**.

The algorithm implies the following properties.

Proposition 2.5.7. *If d is small enough in $\mathbb{Z}$ then there exist exactly $2r$ cylinders of (non-necessarily distinct) orders $k_1, \cdots, k_{2r}$, respectively in $NIZ_{k_1}, \cdots, NIZ_{k_{2r}}$ on which $\Delta^{(r)}$ takes the value d.*

Remark 10. *The "small enough" assumption is precised in the proof.*

It leads to the corollary 2.1.1 we state again here.

Corollary. *For d small enough in $\mathbb{Z}$, we have the formula*

$$\mu^{(r)}(d-1) = \mu^{(r)}(d) \times \frac{1}{\varphi^2}.$$

Remark 11. *The analogous relation in the integer base b case is*

$$\mu^{(r)}(d-(b-1)) = \mu^{(r)}(d) \times \frac{1}{b}.$$

This result is actually intuitive. We focus on the integer base 10 case for instance. The variation of the sum of digits when adding r is actually linked to the creation of carries during the addition : the variation is d if and only if the addition with r creates a precise number of carries (for more details, see [6] Equation (12) page 9). Now, imagine r has, for instance, 2 digits in its expansion and we know the cylinders that creates 4 carries when we add r to our of their elements. To create 5 carries, we just take the same cylinders and concatenate a digit 9 at the right of the leftmost digit of its name : the concatenation divides the measure of the cylinder by b. For instance, if $r = 17$ in base 10, the cylinder C_{69983} is a cylinder our integer-base algorithm will compute and elements in that cylinder create 4 carries. The cylinder C_{699983} is a cylinder that appears at the next order in the New Information Zone and elements of this cylinder creates now 5 carries.

Proof of Proposition 2.5.7. For $k \geq \ell + 2$, there are r cylinders of order k that compose NIZ_k so $\Delta^{(r)}$ takes, at most, r different values on those cylinders. For $k \geq \ell + 2$, let $(d_i^{(k)})_{1 \leq i \leq r}$ be the collection of values taken by $\Delta^{(r)}$ on the cylinders of order k in NIZ_k, repeated with multiplicity. Denote

$$m := min\left\{ d_i^{(k)} \mid i = 1, \cdots, r \text{ and } k = \ell + 2, \ell + 3 \right\}.$$

Now, let $d \leq m$. We observe that due to Lemma 2.5.3 and Lemma 2.5.5, d is actually smaller than any value taken by $\Delta^{(r)}$ on NIZ_ℓ or $\text{NIZ}_{\ell+1}$. Furthermore, for any $i \in [1, r]$ and any $k \in \{\ell+2, \ell+3\}$, there exists a unique $j_i^{(k)} \in \mathbb{N}$ such that $d = d_i^{(k)} - j_i^{(k)}$. Due to Corollary 2.5.6, d will be a value reached by $\Delta^{(r)}$ on the corresponding cylinder of order $k+2j_i^{(k)}$ (by repeated use of Lemma 2.41 and 2.5.5). There are $2r$ cylinders at order $\ell + 2$ and $\ell + 3$ so d will appear exactly on $2r$ cylinders. □

Proof of Corollary 2.1.1. With the same notation as in the proof of Proposition 2.5.7, let $d \leq m$. Then d is a value taken by exactly $2r$ different cylinders of some orders. But, due to Proposition 2.5.7 and Corollary 2.5.6, $\Delta^{(r)}$ will take the value $d - 1$ on the same number of cylinders but with the orders of those cylinders shifted by $+2$ so their measures are divided by φ^2. □

2.6 How to prove $\mu^{(F_\ell)} = \mu^{(1)}$

We propose here two proofs of Theorem 2.4.

There is an analogous relation in the integer base b case. The law of adding b^ℓ is the same as the one of adding 1. It is trivial in base b since the addition $x + b^\ell$ will not change the first digits of x, the addition really consists in adding 1 from a certain position. In the Zeckendorf representation

case, it is not trivial since carries propagate in both directions. Here we propose two proofs : the first one consists into looking at our algorithm described in Section 2.5 in the special case $r = F_\ell$. The other proof relies on the renewal properties of $\mathbb{P}$ we state in Proposition 2.3.2. Before starting, we need to compute $\mu^{(1)}$.

Proposition 2.6.1. *For $d \in \mathbb{Z}$, we have*

$$\mu^{(1)}(d) = \begin{cases} 0 & \text{if } d \geq 2 \\ \frac{1}{\varphi^2} & \text{if } d = 1 \\ \frac{1}{\varphi^{2-2d}} & \text{otherwise.} \end{cases}$$

Proof. Thanks to the algorithm, we explicit the following partition:

$$\mathbb{X} = C_{00} \bigsqcup_{d \leq 0} \left(C_{0(01)^{1-d}} \bigsqcup C_{00(10)^{1-d}} \right). \tag{2.42}$$

We observe that $\Delta^{(1)}(C_{00}) = 1$ and, for every $d \leq 0$,

$$\Delta^{(1)}(C_{0(01)^{1-d}}) = \Delta^{(1)}(C_{00(10)^{1-d}}) = d.$$

Using Proposition 2.3.1, we also have that $\mathbb{P}(C_{00}) = \frac{1}{\varphi^2}$ and

$$\mathbb{P}(C_{0(01)^{1-d}} \bigsqcup C_{00(10)^{1-d}}) = \frac{1}{\varphi^{3-2d}} + \frac{1}{\varphi^{4-2d}} = \frac{1}{\varphi^{2-2d}}.$$

□

2.6.1 First proof using only the algorithm

Let $\ell \geq 3$. We follow the path given by the algorithm. We can rewrite Proposition 2.5.1 in our special case $r = F_\ell$:

1. if $2 \leq k \leq \ell - 1$, $\mathrm{NIZ}_k = \emptyset$,
2. NIZ_ℓ is the union of the $F_{\ell-1}$ first levels of the large tower of order ℓ and
3. if $k \geq \ell + 1$, NIZ_k is the union of the F_ℓ levels between the $F_k - F_\ell^{th}$ and the F_k^{th} levels of the large tower of order k.

Executing the algorithm, we want to compute $\Delta^{(F_\ell)}$ on the levels of NIZ_ℓ. We obtain the following lemma.

Lemma 2.6.2.

$$\Delta^{(F_\ell)}(NIZ_\ell) = 1.$$

Proof. Indeed, consider a cylinder of order ℓ in NIZ_ℓ. Its name has for leftmost digits 000 since the cylinders contains one element of the integer interval $[0, F_{\ell-1} - 1]$. So the addition with F_ℓ considered in these cylinders are

$$\begin{array}{ccccccccc} (x) & & \cdots & 0 & 0 & 0 & \cdots \\ (F_\ell) & + & & & 1 & & \\ \hline (x + F_\ell) & = & \cdots & 0 & 1 & 0 & \cdots \end{array}$$

Thus the variation of the sum of digits is 1. □

We now look at $\mathrm{NIZ}_{\ell+1}$ which is composed of F_ℓ cylinders of order $\ell+1$. We have the following proposition.

Proposition 2.6.3. *In $NIZ_{\ell+1}$, there are*

- $F_{\ell-1}$ *levels on which* $\Delta^{(F_\ell)}=0$ *and*
- $F_{\ell-2}$ *levels on which* $\Delta^{(F_\ell)}=1$.

For technical reasons, we need to separate the proof in two parts : one when ℓ is odd and the other if ℓ is even. However, since both are treated the same way, the difference are only technicalities and we only consider the case where ℓ is odd. We are going to prove that the values taken by $\Delta^{(F_\ell)}$ are the F_ℓ first terms of the sequence

- $1,0,1,0,0,1,1,1,0,0,0,0,0,1,1,1,1,1,1,1,1,\cdots$ (if ℓ is odd)
- $0,1,0,1,1,0,0,0,1,1,1,1,1,0,0,0,0,0,0,0,0,\cdots$ (if ℓ is even) or

where both sequences starts with 0 or 1 depending on the parity and then, the construction consists in concatenating alternatively 0's or 1's a Fibonacci number of times.

Proof of Proposition 2.6.3 if ℓ is odd. Let $\ell'\geq 1$ and $\ell:=2\ell'+1\geq 3$. The highest level in $\mathrm{NIZ}_{\ell+1}$ is the cylinder $C_{00(10)^{\ell'}}$ since it contains $F_{\ell+1}-1$; the lowest is the cylinder $C_{00010^{\ell-3}}$ since it contains $F_{\ell-1}$.

The cylinder of order $\ell+1$ just below $C_{00(10)^{\ell'}}$ is $C_{00(10)^{\ell'-1}01}$. We observe that, if $\ell=3$ (ie $\ell'=1$ or $F_\ell=2$), this last cylinder $C_{00(10)^{\ell'-1}01}$ is actually the same as $C_{00010^{\ell-3}}$. In general, there are two possibilities.

- Either there is no other cylinders of order $\ell+1$ which means $\ell=3$ (ie $\ell'=1$ or $F_\ell=2$). So far, we have mentioned 2 cylinders of order 4 in NIZ_4 and the values of $\Delta^{(F_3)}$ on these cylinders are 0 and 1 (with an easy computation): it is the conclusion of the proposition if $\ell=3$ (if we agree $F_0:=1$).
- Or there are other cylinders to find in $\mathrm{NIZ}_{\ell+1}$ which means $F_\ell>2$ (ie $\ell\geq 5$ or $F_\ell\geq 5$) and implies that there are $F_\ell-F_2\geq 3$ levels to identify.

To continue, we suppose $\ell\geq 5$ (ie $\ell'\geq 2$) and identify the three cylinders that are below $C_{00(10)^{\ell'-1}01}$: they are

- $C_{00(10)^{\ell'-1}00}$ on which $\Delta^{(F_\ell)}$ equals 1,
- $C_{00(10)^{\ell'-2}0101}$ on which $\Delta^{(F_\ell)}$ equals 0 and
- $C_{00(10)^{\ell'-2}0100}$ on which $\Delta^{(F_\ell)}$ equals 0.

Once again, we observe that, if $\ell=5$ (ie $\ell'=2$), $C_{00(10)^{\ell'-2}0100}=C_{00010^{\ell-3}}$. There are again two possibilities.

- Either there is no other cylinders of order $\ell+1$ which means $\ell=5$ (ie $\ell'=2$ or $F_\ell=5$). So far again, we have mentioned 5 cylinders of order 6 in NIZ_6 and $\Delta^{(F_5)}$ takes thrice the value 0 and twice the value 1 on these cylinders : it is the conclusion of the proposition if $\ell=5$.

- Or there are other cylinders to find in $\mathrm{NIZ}_{\ell+1}$ which means $F_\ell > 5$ (ie $\ell \geq 7$ or $F_\ell \geq 13$) and implies that there are $F_\ell - F_5 \geq 8$ levels to identify.

We continue the same way. If we assume there are, for some $k \in [1, \ell-1]$, $F_\ell - F_{2k-1} \geq F_{2k}$ cylinders. There are two kind of that cylinders.

- Those whose names are $C_{00(10)^{\ell'-k+1}00n_{2k-3}\cdots n_2}$ for all $(n_{2k-3}\cdots n_2) \in \mathbb{X}_f$: there are F_{2k-2} such cylinders and one can compute that $\Delta^{(F_\ell)}$ is 1.
- Those whose names are $C_{00(10)^{\ell'-k}010n_{2k-2}\cdots n_2}$ for all $(n_{2k-2}\cdots n_2) \in \mathbb{X}_f$: there are F_{2k-1} such cylinders and one can compute that $\Delta^{(F_\ell)}$ is 0.

We observe that if $\ell = 2k+1$ (ie $\ell' = k$) the last cylinder identified is $C_{00(10)^{\ell'-k}010^{2k-2}}$ and is actually the same as $C_{00010^{\ell-3}}$. So far, we identified $1 + F_1 + (F_2+F_3) + (F_4+F_5) + \cdots + (F_{2k-2}+F_{2k-1}) = F_{2k+1}$ cylinders and

- on $1 + F_2 + F_4 + F_{2k-2} = F_{2k-1}$ of them, $\Delta^{(F_\ell)}$ is 1
- on $F_1 + F_3 + F_5 + \cdots + F_{2k-1} = F_{2k}$ of them, $\Delta^{(F_\ell)}$ is 0

It is exactly the conclusion when $\ell = 2k+1$. □

Finally, we look at $\mathrm{NIZ}_{\ell+2}$ which contains also F_ℓ cylinders of order $\ell+2$.

Proposition 2.6.4.

$$\Delta^{(F_\ell)}(NIZ_{\ell+2}) = 0.$$

Proof. Consider a cylinder of order $\ell+2$ in $\mathrm{NIZ}_{\ell+2}$. Since it contains exactly one integer of the interval $[F_{\ell+1}, F_{\ell+2}-1]$, its name is $C_{0010n_{\ell-1}\cdots n_2}$ for every $(n_{\ell-1}\cdots n_2) \in \mathbb{X}_f$. We compute that, on such a cylinder, $\Delta^{(F_\ell)}$ is null. □

We do not need to prove anything for $\mathrm{NIZ}_{\ell+3}$, the algorithm teach us that $\Delta^{(F_\ell)}$ is going to take the same values, but shifted by -1, on the same number of levels as in $\mathrm{NIZ}_{\ell+1}$. Now it is time to prove our theorem.

Theorem. *For any $\ell \geq 2$, $\mu^{(F_\ell)} = \mu^{(1)}$.*

Proof. We recall

$$\mu^{(1)}(d) = \begin{cases} 0 & \text{if } d \geq 2 \\ \frac{1}{\varphi^2} & \text{if } d = 1 \\ \frac{1}{\varphi^{2-2d}} & \text{otherwise.} \end{cases}$$

For $\ell = 2$, it is trivial. Let $\ell \geq 3$ and $d \in \mathbb{Z}$. It is also trivial when $d \geq 2$ since $\Delta^{(F_\ell)}$ does not take this value. Suppose that $d = 1$, this value appears in $F_{\ell-1}$ levels in NIZ_ℓ, $F_{\ell-2}$ levels in $\mathrm{NIZ}_{\ell+1}$ and nowhere else. So

$$\mu^{(F_\ell)}(1) = \frac{F_{\ell-1}}{\varphi^\ell} + \frac{F_{\ell-2}}{\varphi^{\ell+1}} = \frac{\varphi^{\ell-1}}{\varphi^{\ell+1}} = \frac{1}{\varphi^2} = \mu^{(1)}(1).$$

If $d \leq 0$, the value d will appear, due to Corollary 2.5.6, on $F_{\ell-1}$ levels at order $\ell+1-2d$, F_ℓ levels at order $\ell+2-2d$, $F_{\ell-2}$ levels at order $\ell+3-2d$ and nowhere else. Thus

$$\begin{aligned}\mu^{(F_\ell)}(1) &= \frac{F_{\ell-1}}{\varphi^{\ell+1-2d}} + \frac{F_\ell}{\varphi^{\ell+2-2d}} + \frac{F_{\ell-2}}{\varphi^{\ell+3-2d}} \\ &= \frac{\varphi^{\ell+1}}{\varphi^{\ell+3-2d}} \\ &= \frac{1}{\varphi^{2-2d}} \\ &= \mu^{(1)}(d).\end{aligned}$$

□

2.6.2 Another proof using also the renewal

This second proof uses the renewal properties of the measure $\mathbb{P}$ described in Subsection 2.3.2. Let $\ell \geq 3$ and $x \in \mathbb{X}$. We consider the addition

$$\begin{array}{cccccccc} (x) & & \cdots & x_{\ell+1} & x_\ell & x_{\ell-1} & \cdots & x_2 \\ (F_\ell) & + & & & 1 & 0 & \cdots & 0 \end{array}$$

We are going to discuss on 3 possibilities : $x_{\ell-1} = 1$ or $x_\ell = x_{\ell-1} = 0$ or $x_\ell = 1$ and we will conclude using the law of total probability.

1. If $x_{\ell-1} = 1$, we are considering the addition

$$\begin{array}{cccccccccc} (x) & & \cdots & x_{\ell+1} & 0 & 1 & 0 & x_{\ell-3} & \cdots & x_2 \\ (F_\ell) & + & & & 1 & 0 & 0 & 0 & \cdots & 0 \end{array}$$

We give an expression that transfer the dependence on ℓ on x through the following lemma.

Lemma 2.6.5. *For almost every $x \in \mathbb{X}$ (for the measure $\mathbb{P}$ conditioned by $x_{\ell-1} = 1$)*

$$\Delta^{(F_\ell)}(x) = \Delta^{(F_3)}(\sigma^\ell x) - 1.$$

Proof. Let $x \in \mathbb{X}$. Suppose x has two consecutive 0's at indices $\geq \ell+1$. This assumption is of probability one. There are two cases.

(a) Either there exists $k > 0$ such that $x = \overline{\cdots 0(01)^k 010 x_{\ell-3} x_2}$,

(b) or there exists $k > 0$ such that $x = \overline{\cdots 00(10)^k 010 x_{\ell-3} x_2}$.

The rest of the proof is just computations we made in Subsection 2.2.1. □

Thus, it follows the formula

$$\mathbb{P}\left(\Delta^{(F_\ell)}(x) = d \mid x_{\ell-1} = 1\right) = \mathbb{P}\left(\Delta^{(F_3)}(\sigma^\ell x) = d+1 \mid x_{\ell-1} = 1\right).$$

But, due to the Proposition 2.3.2, the law of distribution of $\sigma^\ell x$ conditioned by the event $x_{\ell-1} = 1$ is the same as the law of distribution of x conditioned by the event $x_2 = 0$. We write it shortly as follows

$$\mathcal{L}\left(\sigma^\ell x \mid x_{\ell-1} = 1\right) = \mathcal{L}\left(x \mid x_2 = 0\right).$$

It implies

$$\mathbb{P}\left(\Delta^{(F_\ell)}(x) = d \mid x_{\ell-1} = 1\right) = \mathbb{P}\left(\Delta^{(F_3)}(x) = d+1 \mid x_2 = 0\right). \tag{2.43}$$

2. If $x_\ell = x_{\ell-1} = 0$, we now consider the addition

$$\begin{array}{cccccccc} (x) & & x_{\ell+1} & 0 & 0 & x_{\ell-2} & \cdots & x_2 \\ (F_\ell) & + & & 1 & 0 & 0 & \cdots & 0 \end{array}$$

We proceed as before by he following lemma that transfer the dependence on ℓ on x

Lemma 2.6.6. *For almost every $x \in \mathbb{X}$ (for the measure $\mathbb{P}$ conditioned by $x_\ell = x_{\ell-1} = 0$)*

$$\Delta^{(F_\ell)}(x) = \Delta^{(F_2)}(\sigma^\ell x).$$

Proof. The $\ell - 1$ first digits of $x + F_\ell$ are the same as those of x. The addition will modify only digits of indices $\geq \ell$. Thus, this addition only consists into adding 1 to $\sigma^\ell x$. □

Thus, it follows the formula

$$\mathbb{P}\left(\Delta^{(F_\ell)}(x) = d \mid x_\ell = x_{\ell-1} = 0\right) = \mathbb{P}\left(\Delta^{(F_3)}(\sigma^\ell x) = d \mid x_\ell = x_{\ell-1} = 0\right).$$

The renewal properties of $\mathbb{P}$ gives that

$$\mathcal{L}\left(\sigma^\ell x \mid x_\ell = x_{\ell-1} = 0\right) = \mathcal{L}\left(x \mid x_2 = 0\right).$$

It implies

$$\mathbb{P}\left(\Delta^{(F_\ell)}(x) = d \mid x_\ell = x_{\ell-1} = 0\right) = \mathbb{P}\left(\Delta^{(F_2)}(x) = d \mid x_2 = 0\right). \tag{2.44}$$

3. If $x_\ell = 1$, we are considering the following addition

$$\begin{array}{ccccccccccc} (x) & & \cdots & x_{\ell+2} & 0 & 1 & 0 & x_{\ell-2} & \cdots & x_2 \\ (F_\ell) & + & & & & 1 & 0 & 0 & \cdots & 0 \\ \hline (x+F_\ell) & = & \cdots & x_{\ell+2} & 0 & 2 & 0 & x_{\ell-2} & \cdots & x_2 \\ & & & & +1 & -2 & & +1 & & \end{array}$$

Since it is the first time the addition creates carries in both directions, arguments are getting difficult. However, the good expansion of $(x + r)$ for the digits of indices $\geq \ell + 1$ is given by

the expansion of $\sigma^{\ell+1}x+1$. Also, the expansion of the digits of indices $\leq \ell$ is given by the expansion of $\overline{x_{\ell-2}\cdots x_2}+F_{\ell-2}$. In other words, for every $k\geq 2$

$$\begin{aligned}\Delta_k^{F_\ell}(x) &= s_k(x+F_\ell)-s_k(x)\\ &= s_k(\sigma^{\ell+1}x+1)+s_k(F_{\ell-2}+\overline{x_{\ell-2}\cdots x_2})\\ &\quad -(s_k(\sigma^{\ell+1}x)+1+s_k(\overline{x_{\ell-2}\cdots x_2}))\\ &= \Delta_k^{(1)}(\sigma^{\ell+1}x)+\Delta_k^{(F_{\ell-2})}(\overline{x_{\ell-2}\cdots x_2})-1.\end{aligned}$$

Thus, for almost every $x\in\mathbb{X}$ (for the measure $\mathbb{P}$ conditioned by $x_\ell=1$)

$$\Delta^{(F_\ell)}(x)=\Delta^{(1)}(\sigma^{\ell+1}x)+\Delta^{(F_{\ell-2})}(\overline{x_{\ell-2}\cdots x_2})-1. \tag{2.45}$$

For the term $\Delta^{(F_{\ell-2})}(\overline{x_{\ell-2}\cdots x_2})$, we can take any combinations $(x_{\ell-2},\cdots,x_2)\in\mathbb{X}_f$: the almost every $x\in\mathbb{X}$ assumption depends on digits of indices $\geq \ell+1$ as we see in (2.45). We have the following lemma.

Lemma 2.6.7. *For any* $(x_{\ell-2},\cdots,x_2)\in\mathbb{X}_f$

$$\Delta^{(F_{\ell-2})}(\overline{x_{\ell-2}\cdots x_2})=\begin{cases}1 & \text{for } F_{\ell-2} \text{ choices}\\ 0 & \text{for } F_{\ell-3} \text{ choices.}\end{cases}$$

Proof. Let $(x_{\ell-2},\cdots,x_2)\in\mathbb{X}_f$. The integer $\overline{x_{\ell-2}\cdots x_2}\in[0,F_{\ell-1}-1]$. We have the partition

$$[0,F_{\ell-1}-1]=[0,F_{\ell-3}-1]\sqcup[F_{\ell-3},F_{\ell-1}-1].$$

The integer interval $[0,F_{\ell-3}-1]$ is part of $\mathrm{NIZ}_{\ell-2}$ the first non-empty New Information Zone for $r=F_{\ell-2}$. Thus, due to Lemma 2.6.2, $\Delta^{(F_{\ell-2})}=0$ on that part. For the other integer interval $[F_{\ell-3},F_{\ell-1}-1]$, we observe that it is part of $\mathrm{NIZ}_{\ell-1}$ the first non-empty New Information Zone for $r=F_{\ell-2}$. There are as many integers as cylinders of order $\ell-1$ in that integer interval and to each integer corresponds a unique cylinders. Then, it follows from Proposition 2.6.3, that $F_{\ell-4}$ of these integers are mapped to 1 by $\Delta^{(F_{\ell-2})}$ and $F_{\ell-3}$ are mapped to 0. To sum up, $F_{\ell-3}$ combinations $(x_{\ell-2},\cdots,x_2)\in\mathbb{X}_f$ gives 0 by $\Delta^{(F_{\ell-2})}$ and $F_{\ell-2}$ combinations gives 1.

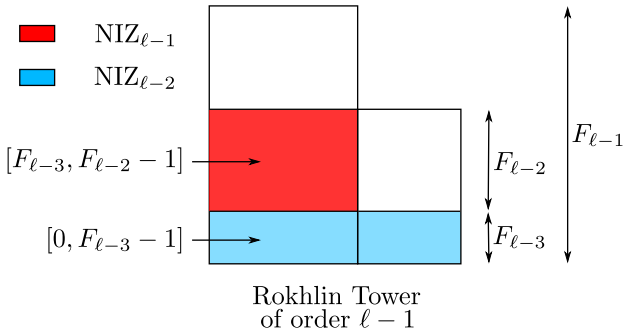


Figure 2.12: $\mathrm{NIZ}_{\ell-1}$ and $\mathrm{NIZ}_{\ell-2}$ in the Rokhlin towers of order $\ell-1$.

□

Combining Lemma 2.6.7 and Equation (2.45), we obtain that for almost every $x \in \mathbb{X}$ (for the measure $\mathbb{P}$ conditioned by $x_\ell = 1$)

$$\Delta^{(F_\ell)}(x) = \begin{cases} \Delta^{(1)}(\sigma^{\ell+1}x) & \text{for } F_{\ell-2} \text{ choices} \\ \Delta^{(1)}(\sigma^{\ell+1}x) - 1 & \text{for } F_{\ell-3} \text{ choices.} \end{cases} \tag{2.46}$$

where, by "*choices*", we mean "*choices of* $(x_{\ell-2}, \cdots, x_2) \in \mathbb{X}_f$". For simplicity and the rest of this part, we will mention as

(a) *Condition 1* (or *Cdt 1*) the $F_{\ell-2}$ choices of $(x_{\ell-2}, \cdots, x_2) \in \mathbb{X}_f$ that implies $\Delta^{(F_\ell)}(x) = \Delta^{(1)}(\sigma^{\ell+1}x)$ and

(b) *Condition 2* (or *Cdt 1*) the $F_{\ell-3}$ choices of $(x_{\ell-2}, \cdots, x_2) \in \mathbb{X}_f$ that implies $\Delta^{(F_\ell)}(x) = \Delta^{(1)}(\sigma^{\ell+1}x) - 1$.

Thus, from (2.46), it follows that

$$\begin{cases} \mathbb{P}\left(\Delta^{(F_\ell)}(x) = d \mid x_\ell = 1 \text{ and Cdt } 1\right) & = \mathbb{P}\left(\Delta^{(1)}(\sigma^{\ell+1}x) = d \mid x_\ell = 1 \text{ and Cdt } 1\right) \\ \mathbb{P}\left(\Delta^{(F_\ell)}(x) = d \mid x_\ell = 1 \text{ and Cdt } 2\right) & = \mathbb{P}\left(\Delta^{(1)}(\sigma^{\ell+1}x) = d+1 \mid x_\ell = 1 \text{ and Cdt } 2\right). \end{cases}$$

The renewal properties of $\mathbb{P}$ give that

$$\mathcal{L}\left(\sigma^{\ell+1}x \mid x_\ell = 1 \text{ and Cdt } 1\right) = \mathcal{L}\left(\sigma^{\ell+1}x \mid x_\ell = 1 \text{ and Cdt } 2\right) = \mathcal{L}\left(x \mid x_2 = 0\right).$$

It implies

$$\mathbb{P}\left(\Delta^{(F_\ell)}(x) = d \mid x_\ell = 1 \text{ and Cdt } 1\right) = \mathbb{P}\left(\Delta^{(1)}(x) = d \mid x_2 = 0\right) \tag{2.47}$$

$$\mathbb{P}\left(\Delta^{(F_\ell)}(x) = d \mid x_\ell = 1 \text{ and Cdt } 2\right) = \mathbb{P}\left(\Delta^{(1)}(x) = d+1 \mid x_2 = 0\right). \tag{2.48}$$

Now, we have all the ingredients to prove Theorem 2.1.2.

Proof of Theorem 2.1.2. From (2.43), (2.44), (2.47), (2.48) and the law of total probability, one find

$$\begin{aligned} \mathbb{P}\left(\Delta^{(F_\ell)}(x) = d\right) &= \mathbb{P}\left(x_{\ell-1} = 1\right)\mathbb{P}\left(\Delta^{(F_\ell)}(x) = d \mid x_{\ell-1} = 1\right) \\ &\quad + \mathbb{P}\left(x_\ell = x_{\ell-1} = 0\right)\mathbb{P}\left(\Delta^{(F_\ell)}(x) = d \mid x_\ell = x_{\ell-1} = 0\right) \\ &\quad + \mathbb{P}\left(x_\ell = 1 \text{ and Cdt } 1\right)\mathbb{P}\left(\Delta^{(F_\ell)}(x) = d \mid x_\ell = 1 \text{ and Cdt } 1\right) \\ &\quad + \mathbb{P}\left(x_\ell = 1 \text{ and Cdt } 2\right)\mathbb{P}\left(\Delta^{(F_\ell)}(x) = d \mid x_\ell = 1 \text{ and Cdt } 2\right) \\ &= \frac{F_{\ell-2}}{\varphi^\ell}\ \mathbb{P}\left(\Delta^{(F_3)}(x) = d+1 \mid x_2 = 0\right) \\ &\quad + \frac{F_{\ell-1}}{\varphi^\ell}\ \mathbb{P}\left(\Delta^{(1)}(x) = d \mid x_2 = 0\right) \\ &\quad + \frac{F_{\ell-2}}{\varphi^{\ell+1}}\ \mathbb{P}\left(\Delta^{(1)}(x) = d \mid x_2 = 0\right) \\ &\quad + \frac{F_{\ell-3}}{\varphi^{\ell+1}}\ \mathbb{P}\left(\Delta^{(1)}(x) = d+1 \mid x_2 = 0\right). \end{aligned}$$

We claim that for any $d \in \mathbb{Z}$,

$$\mathbb{P}\left(\Delta^{(1)}(x) = d \mid x_2 = 0\right) = \mathbb{P}\left(\Delta^{(F_3)}(x) = d \mid x_2 = 0\right).$$

Indeed, it is trivial if $d > 1$. For $d \leq 1$, we use the partition (2.42) and (left to the reader) the analogous partition given by the algorithm for $\Delta^{(F_3)}$. We find that, for $d \leq 1$

$$\mathbb{P}\left(\Delta^{(1)}(x) = d \mid x_2 = 0\right) = \frac{\mathbb{P}(C_{00(10)^{1-d}})}{\mathbb{P}(x_2 = 0)} = \frac{1}{\varphi^{3-2d}}$$

as well as

$$\mathbb{P}\left(\Delta^{(F_3)}(x) = d \mid x_2 = 0\right) = \frac{\mathbb{P}(C_{0(01)^{1-d}00} \sqcup C_{00(10)^{2-d}})}{\mathbb{P}(x_2 = 0} = \frac{1}{\varphi^{3-2d}}.$$

Coming back to the law of total probability, we obtain

$$\begin{aligned}
\mathbb{P}\left(\Delta^{(F_\ell)}(x) = d\right) &= \frac{F_{\ell-3} + \varphi F_{\ell-2}}{\varphi^{\ell+1}} \mathbb{P}\left(\Delta^{(1)}(x) = d+1 \mid x_2 = 0\right) \\
&\quad + \frac{F_{\ell-2} + \varphi F_{\ell-1}}{\varphi^{\ell+1}} \mathbb{P}\left(\Delta^{(1)}(x) = d \mid x_2 = 0\right) \\
&= \frac{1}{\varphi^2} \mathbb{P}\left(\Delta^{(1)}(x) = d+1 \mid x_2 = 0\right) \\
&\quad + \frac{1}{\varphi} \mathbb{P}\left(\Delta^{(1)}(x) = d \mid x_2 = 0\right) \\
&= \mathbb{P}\left(\Delta^{(1)}(x) = d\right).
\end{aligned}$$

□

2.7 $\Delta^{(r)}$ as a mixing process

We work on the probability space $(\mathbb{X}, \mathscr{B}(\mathbb{X}), \mathbb{P})$. For a given integer r, $\Delta^{(r)}$ is viewed as a random variable with law $\mu^{(r)}$ by Corollary 2.4.2 (the randomness comes from the argument x of $\Delta^{(r)}$, considered as a random outcome in $\mathbb{X}$ with law $\mathbb{P}$). In this section, we will decompose $\Delta^{(r)}$ as a sum composed of a finite number of random variables satisfying a universal inequality on their mixing coefficients, as in [6].

2.7.1 The process

The case $r = 0$ is irrelevant so we assume $r \geq 1$. For $1 \leq i \leq \rho(r)$, we will write B_i as the i^{th} block present in the expansion of r, starting from the unit digit. We define $r[i]$ as the integer whose expansion is given by the i first blocks of the expansion of r (see Figure below). We observe that $r[\rho(r)] = r$.

$$
\begin{aligned}
r = \quad & \overline{\underbrace{1010}_{B_4}\ 0\ \underbrace{10}_{B_3}\ 0\ \underbrace{1010}_{B_2}\ 000\ \underbrace{101}_{B_1}} \\
r[1] = \quad & \overline{101} \\
r[2] = \quad & \overline{1010\ 000\ 101} \\
r[3] = \quad & \overline{10\ 0\ 1010\ 000\ 101} \\
r[4] = \quad & \overline{1010\ 0\ 10\ 0\ 1010\ 000\ 101}
\end{aligned}
$$

Figure 2.13: Example of the construction of $(r[i])_{i=0,\cdots,\rho(r)}$.

With the convention $r[0] := 0$, we observe the trivial equality

$$r = \sum_{i=1}^{\rho(r)} r[i] - r[i-1].$$

For $1 \leq i \leq \rho(r)$, we define almost everywhere on $\mathbb{X}$ (see Subsection 2.4)

$$X_i^{(r)} := \Delta^{(r[i]-r[i-1])} \circ T^{r[i-1]}.$$

Since $r[i]-r[i-1] = \overline{B_i 0 \cdots 0}$, the function $X_i^{(r)}$ is a random variable corresponding to the action of the i^{th} block B_i once the previous blocks have already been taken into consideration. From (2.32), we get

$$\Delta^{(r)} = \sum_{i=1}^{\rho(r)} X_i^{(r)}.$$

In particular, if $x \in \mathbb{X}$ is randomly chosen with law $\mathbb{P}$, then $\sum_{i=1}^{\rho(r)} X_i^{(r)}(x)$ follows the law $\mu^{(r)}$.

2.7.2 The α-mixing coefficients on the actions of blocks

This part is devoted to the proof of one of the main theorem stated in the Introduction: we are going to show that the α-mixing coefficients for the process $(X_i^{(r)})_{i=1,\cdots,\rho(r)}$ satisfy a universal upper bound which is independent of r. In particular, these coefficients exponentially decrease to 0.

Theorem. *The α-mixing coefficients of $(X_i^{(r)})_{i=1,\cdots,\rho(r)}$ satisfy*

$$\forall k \geq 1, \qquad \alpha(k) \leq 12 \left(1 - \frac{1}{\varphi^8}\right)^{\frac{k}{6}} + \frac{1}{\varphi^{2k}}.$$

Proof. Let $k, p \geq 1$. Let $A \in \sigma(X_i^{(r)} : 1 \leq i \leq p)$ and $B \in \sigma(X_i^{(r)} : i \geq k+p)$. Let $C \in \sigma(X_i^{(r)} : p < i < k+p)$ such that $\mathbb{P}(C) > 0$. Then, we have the following inequality

$$\begin{aligned} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \quad &\leq |\mathbb{P}(A \cap B) - \mathbb{P}_C(A \cap B)| \\ &\quad + |\mathbb{P}_C(A \cap B) - \mathbb{P}_C(A)\mathbb{P}_C(B)| \\ &\quad + |\mathbb{P}_C(A)\mathbb{P}_C(B) - \mathbb{P}(A)\mathbb{P}(B)|. \end{aligned} \tag{2.49}$$

Using (2.25), we get

$$|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \leq 3\mathbb{P}(\overline{C}) + |\mathbb{P}_C(A \cap B) - \mathbb{P}_C(A)\mathbb{P}_C(B)|. \tag{2.50}$$

As in [6] (Lemma 4.3), we will consider an event C which has a probability close to 1 and such that, conditionally to C, the events A and B are "almost" independent. A difference with [6] is that, as the digits of a random Zeckendorf-adic numbers are not independent, the last term at the right-hand side of (2.50) will totally vanish and, thus, some extra work is needed. However, we can anticipate that this problem of "almost" independence will be solved using the inequality on the ϕ-mixing coefficients (on the law of coordinates) stated in Proposition 2.3.5.

To define the event C, we use the ideas developed in Subsection 2.2.3 (especially Corollaries 2.2.7 and 2.2.8). Since we are working with blocks, we introduce ℓ_i as the number of patterns 10 in B_i, the i^{th} block of r. We also introduce n_i as the minimal index of digits in B_i. Then, we define, for $i = 1, \cdots, \rho(r)$, the set $\mathrm{Adm}(i)$ as the set of indices corresponding to the digits of x involved in the assumptions of Corollaries 2.2.7 and 2.2.8 if we want to apply them in the area of the block B_i in r. More precisely, if $\ell_i = 1$, then

$$\mathrm{Adm}(i) := [n_i - 2, n_i + 3].$$

And, if $\ell_i \geq 2$ then

$$\mathrm{Adm}(i) := [n_i - 2, n_i + 1] \sqcup [n_i - 2 + 2\ell_i, n_i + 1 + 2\ell_i].$$

Observe that $|\mathrm{Adm}(i)| \leq 8$.

$$r = \overline{\cdots\ 00100\ \cdots} \qquad r = \overline{\cdots\ 0010(10)^{\ell(i)-2}100\ \cdots}$$

$$x = \overline{\cdots\ 000000\cdots} \qquad x = \overline{\cdots\ 0000\quad \cdots \quad 0000\cdots}$$

Figure 2.14: Indices of $\mathrm{Adm}(i)$.

Then, we define for $i = 1, \cdots, \rho(r)$

$$C^{(i)} := \{x \in \mathbb{X} \ : \ \forall j \in \mathrm{Adm}(i),\ x_j = 0\}.$$

We also define the events

$$C_1 := \bigcup_{\substack{i=p+2 \\ \text{even}}}^{p+\frac{k}{3}} C^{(i)} \qquad \text{and} \qquad C_2 := \bigcup_{\substack{i=p+\frac{2k}{3} \\ \text{even}}}^{p+k-2} C^{(i)}$$

and, finally

$$C := C_1 \cap C_2.$$

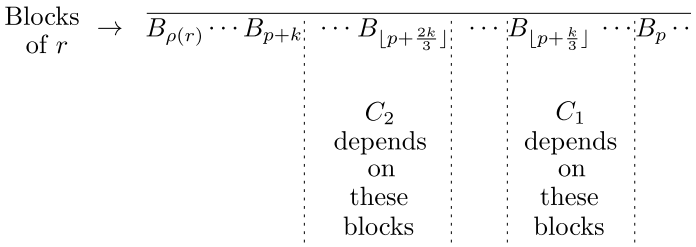


Figure 2.15: Location of the events C_1 and C_2 (dots represent blocks or possible zeros).

Thus, x belongs to C if and only if the expansion of x satisfies the hypotheses of Corollary 2.2.7 or 2.2.8 for, at least, two blocks placed on the leftmost third of the window (for at least one block) and the rightmost third of the window (for at least one block).

We prove now that C is an event which has a high probability to happen. We claim that

$$\mathbb{P}(C) \geq 1 - 2(1 - \frac{1}{\varphi^8})^{\frac{k}{6}}. \tag{2.51}$$

Indeed, denoting $\overline{C_1}$ (resp. $\overline{C_2}$) the complement of C_1 (resp. C_2) in $\mathbb{X}$, we have

$$\mathbb{P}(C) = 1 - \mathbb{P}(\overline{C_1}) - \mathbb{P}(\overline{C_2}) + \mathbb{P}(\overline{C_1} \cap \overline{C_2}) \geq \mathbb{P}(C_1) + \mathbb{P}(C_2) - 1.$$

Then, we obtain from Lemma 2.3.7

$$\begin{aligned}
\mathbb{P}(C_1) &= 1 - \mathbb{P}\left(\bigcap_{\substack{i=p+2 \\ \text{even}}}^{p+\frac{k}{3}} \overline{C^{(i)}} \right) \\
&= 1 - \prod_{\substack{i=p+2 \\ \text{even}}}^{p+\frac{k}{3}} \mathbb{P}\left(\overline{C^{(i)}} \;\middle|\; \bigcap_{\substack{j=p+2 \\ \text{even}}}^{i-2} \overline{C^{(j)}} \right) \\
&\geq 1 - (1 - \frac{1}{\varphi^8})^{\frac{k}{6}}
\end{aligned}$$

because the product contains $\lfloor \frac{p+\frac{k}{3}-p+2-2}{2} \rfloor = \lfloor \frac{k}{6} \rfloor$ terms. We show that $\mathbb{P}(C_2)$ satisfies the same inequality. We thus obtain (2.51). Observe that, if $k \geq 194$, $\mathbb{P}(C) > 0$.

Now, we want to control the term $|\mathbb{P}_C(A \cap B) - \mathbb{P}_C(A)\mathbb{P}_C(B)|$ that appears in (2.50). We want to use Proposition 2.3.5, and for that we have to clarify which digits of x the events A and B depends on.

But, conditionally to C and thanks to Corollaries 2.2.7 and 2.2.8, we have that the actions of the p first blocks will only modify digits of indices $\leq N_1 + 2$ where N_1 is the index of digit of the leftmost 1 of the block $B_{\lfloor p+\frac{k}{3} \rfloor}$ of r. Thus, there exists $A' \in \sigma(x_i \ : \ i \leq N_1 + 2)$ such that $A \cap C = A' \cap C$. Likewise, the actions of the $\rho(r) - k - p + 1$ last blocks will only modify digits of indices $\geq N_2$ where N_2 the index of the rightmost 0 of the block $B_{\lfloor p+\frac{2k}{3} \rfloor}$. Thus, there also exists $B' \in \sigma(x_i \ : \ i \geq N_2)$ such that $B' \cap C = B \cap C$.

$$r = \cdots 00\overbrace{(10)^{\ell-1}100}^{B_{\lfloor p+\frac{2k}{3}\rfloor}} \cdots 00\overbrace{10(10)^{\ell'-1}0}^{B_{\lfloor p+\frac{k}{3}\rfloor}} \cdots$$
$$\uparrow r_{N_2} \qquad \uparrow r_{N_1}$$

For simplicity, $\ell := \ell(\lfloor p + \frac{2k}{3} \rfloor)$.
$\ell' := \ell(\lfloor p + \frac{k}{3} \rfloor)$

Figure 2.16: Location of indices N_1 and N_2 in the expansion of r.

It follows

$$|\mathbb{P}_C(A \cap B) - \mathbb{P}_C(A)\mathbb{P}_C(B)| = |\mathbb{P}_C(A' \cap B') - \mathbb{P}_C(A')\mathbb{P}_C(B')| \tag{2.52}$$

Then, we get

$$\begin{aligned} |\mathbb{P}_C(A' \cap B') - \mathbb{P}_C(A')\mathbb{P}_C(B')| \leq\ & |\mathbb{P}_C(A' \cap B') - \mathbb{P}(A' \cap B')| \\ & + |\mathbb{P}(A' \cap B') - \mathbb{P}(A')\mathbb{P}(B')| \\ & + |\mathbb{P}(A')\mathbb{P}(B') - \mathbb{P}_C(A')\mathbb{P}_C(B')|. \end{aligned}$$

The term $|\mathbb{P}(A' \cap B') - \mathbb{P}(A')\mathbb{P}(B')|$ can be controlled using the ϕ-mixing coefficient on the law of coordinates. Indeed, observe that $N_2 - N_1$ is at least about order k. Then, using (2.25) and Propositions 2.3.4 and 2.3.5, we obtain

$$|\mathbb{P}_C(A' \cap B') - \mathbb{P}_C(A')\mathbb{P}_C(B')| \leq 3\mathbb{P}(\overline{C}) + \frac{1}{2}\phi(k) \tag{2.53}$$

Finally, combining (2.50), (2.52) and (2.53), we obtain

$$|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \leq 6\mathbb{P}(\overline{C}) + \frac{1}{2}\phi(k).$$

Taking the supremum on A and B, we conclude the proof. □

Questions ouvertes et perspectives

Cette thèse a permis d'apporter un nouveau point de vue au problème de variation de la somme des chiffres quand on ajoute un entier fixé. Cette approche, utilisant des techniques probabilistes et de théorie ergodique, se révèle être plutôt fructueuse. On nuancera le propos en précisant que cette approche n'est pas meilleure que l'approche arithméticienne de l'état de l'art mais lui est complémentaire : chacune de ses approches possède ses propres qualités et ses propres défauts.

Au sortir de la thèse, des questions restent en suspens, des perspectives s'ouvrent. Nous en dressons ici une liste non-exhaustive.

En base entière $b \geq 2$

Dans ce système de représentation, nous avons établi un TCL avec vitesse. L'optimalité de la vitesse est une première question que l'on peut se poser. D'abord, par la relation (1.3), on a obtenu une vitesse de convergence des moments partiels de la mesure $\mu^{(r)}$ renormalisée vers ceux d'une loi normale en $\sqrt{\rho(r)}$. C'est une vitesse très satisfaisante : elle est équivalente à celle que l'on obtient dans le cas d'une somme de $\rho(r)$ variables aléatoires iid. Or, c'est à peu près avec ce même nombre de variables que l'on décompose $\Delta^{(r)}$. Néanmoins, cette vitesse ralentit quand on souhaite montrer la convergence des fonctions de répartitions : elle devient alors de l'ordre de $\rho(r)^{\frac{1}{8}}$. On se pose donc les deux questions suivantes.

Questions. *Est-ce une vitesse optimale ? Si non, comment l'améliorer ?*

La technique d'approximation employée dans la Sous-section 1.5.2 ne semble pas améliorable. Une piste d'amélioration serait d'avoir une meilleure borne des coefficients de mélange (se référer à la Proposition 1.4.3) voire un meilleur type de mélange. Le contrôle de la variance pourrait, a priori, être une seconde piste de réflexion. Cependant, le Théorème 1.1.2 indique que la variance croît linéairement en $\rho(r)$ quand le nombre de blocs augmente. On peut tout de même se poser des questions quant à l'optimalité des constantes de majoration et de minoration que l'on obtient dans le Théorème 1.1.2. Comparativement aux résultats de Emme et Prikhod'ko dans [12] ainsi qu'à ceux de Spiegelhofer et Wallner dans [28], il semble assez clair que nos constantes ne sont pas optimales. On nuancera le propos en précisant que les auteurs pré-cités travaillaient en base 2. Or, le système binaire semble avoir un comportement spécifique par rapport aux autres bases entières $b \geq 3$. Il semble judicieux de séparer le cas $b = 2$ du cas $b \geq 3$. Deux nouvelles questions s'offrent alors à nous.

Questions. *Quelles sont les constantes de majoration et de minoration optimales pour* $\mathrm{Var}(\mu^{(r)})$ *? De plus, à un nombre fixé de blocs, quels sont les entiers réalisant la majoration (ou la minoration) de la variance ?*

Ce problème du contrôle de variance semble être le point le plus délicat dans le travail, quel que soit le point de vue adopté sur le problème.

En représentation de Zeckendorf

Ce système de représentation est bien plus difficile à appréhender que le système en base entière. Ainsi, une formule de récurrence similaire à celle (quasiment triviale) en base entière (1) semble être difficile à établir. Or cette formule permettait d'avoir notre contrôle de variance. D'où la question suivante.

Question. *Existe-t-il une formule semblable à* (1) *valable dans ce système ?*

On peut également penser à une autre piste pour contrôler cette variance. Nous sommes dans le cas où $\Delta^{(r)}$ est la somme de variables aléatoires $X_i^{(r)}$ formant un processus mélangeant, donc, par définition, proche d'un processus indépendant. Contrôler la variance d'une somme de variables indépendantes est "facile", il suffit de contrôler chaque terme. Dans notre cas, on peut donc penser pouvoir obtenir une majoration et une minoration de la variance de $\Delta^{(r)}$ en fonction des coefficients de mélange. Néanmoins, il semble déraisonnable de ne pas, d'abord, introduire de "bonnes hypothèses" vérifiées par les variables $X_i^{(r)}$, $i = 1, \cdots, \rho(r)$. On se pose donc les questions plus générales suivantes.

Questions. *Est-il possible d'obtenir un contrôle de la variance d'une somme de variables aléatoires mélangeantes en fonction des coefficients de mélange ? Si non, peut-on tout de même en trouver un, quitte à introduire des hypothèses sur les variables (valable dans notre problème) ?*

Bien que nous n'ayons pas de formule similaire à (1), nous avons réussi à obtenir un algorithme (adaptable en base entière) permettant le calcul de $\mu^{(r)}$. Cet algorithme a permis de déduire, en particulier, un contrôle de la queue de distribution de $\mu^{(r)}$. En effet, si d est un entier relatif "assez petit" alors $\mu^{(r)}(d)$ vaut une constante indépendante de d divisée par φ^{2d}. On se pose donc la question suivante.

Question. *À partir de quel entier relatif d obtient-on le contrôle de la queue de distribution de* $\mu^{(r)}$ *?*

La réponse proviendra certainement de la démonstration du Corollaire 2.1.1. De plus, l'algorithme étant adaptable en base entière, on peut espérer que la réponse à cette question le soit également. Ainsi, préciser cette hypothèse d'être "assez petit" permettra de mieux connaître la mesure $\mu^{(r)}$ et, en base binaire, de simplifier (légèrement) la Conjecture de Cusick.

D'autres systèmes de numération

La représentation de Zeckendorf est un cas particulier de systèmes de numération d'Ostrowski (se référer à [20]). Pour rappel, ces systèmes découlent du développement en fraction continue d'un irrationnel ($\frac{1}{\varphi}$ pour la représentation de Zeckendorf). Une question naturelle est donc la suivante.

Question. *Peut-on adapter la même méthodologie à un système de numération d'Ostrowski arbitraire ?*

Le cas particulier de la représentation de Zeckendorf s'étant révélé assez relevé, la question semble difficile. On peut penser à une autre forme de généralisation. En effet, derrière la représentation de Zeckendorf se cache la suite de Fibonacci. Une autre suite bien connue permettant de définir un système de numération est la suite de Tribonacci. Cette suite est très liée à la fractale de Rauzy. Peut-être est-il possible de faire apparaître cette fractale en étudiant notre problème de variation de la somme des chiffres.

Question. *Peut-on adapter la même méthodologie au système de numération construit à partir de la suite de Tribonacci (ou de ses extensions) ?*

Bibliographie

[1] G. Barat and P. Liardet. Dynamical systems originated in the Ostrowski alpha-expansion. *Ann. Univ. Sci. Budapest. Sect. Comput.*, 24 :133–184, 2004.

[2] Valérie Berthé. Autour du système de numération d'Ostrowski. volume 8, pages 209–239. 2001. Journées Montoises d'Informatique Théorique (Marne-la-Vallée, 2000).

[3] Jean Bésineau. Indépendance statistique d'ensembles liés à la fonction "somme des chiffres". In *Séminaire de Théorie des Nombres, 1970-1971 (Univ. Bordeaux I, Talence), Exp. No. 21*, page 20 pp. Lab. Théorie des Nombres. 1971.

[4] Patrick Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, third edition, 1995. A Wiley-Interscience Publication.

[5] Richard C. Bradley. Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.*, 2 :107–144, 2005. Update of, and a supplement to, the 1986 original.

[6] Thierry de la Rue, Élise Janvresse, and Yohan Hosten. A central limit theorem for the variation of the sum of digits, 2021.

[7] Paul Doukhan, Pascal Massart, and Emmanuel Rio. The functional central limit theorem for strongly mixing processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 30(1) :63–82, 1994.

[8] Michael Drmota, Manuel Kauers, and Lukas Spiegelhofer. On a conjecture of Cusick concerning the sum of digits of n and $n + t$. *SIAM J. Discrete Math.*, 30(2) :621–649, 2016.

[9] Michael Drmota, Clemens Müllner, and Lukas Spiegelhofer. Primes as sums of fibonacci numbers, 2021.

[10] Jordan Emme and Pascal Hubert. Normal distribution of correlation measures of binary sum-of-digits functions. working paper or preprint, October 2018.

[11] Jordan Emme and Pascal Hubert. Central limit theorem for probability measures defined by sum-of-digits function in base 2. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 19(2) :757–780, 2019.

[12] Jordan Emme and Alexander Prikhod'ko. On the asymptotic behavior of density of sets defined by sum-of-digits function in base 2. *Integers*, 17 :Paper No. A58, 28, 2017.

[13] N. Pytheas Fogg. *Substitutions in dynamics, arithmetics and combinatorics*, volume 1794 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2002. Edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel.

[14] Martin Griffiths. Digit proportions in Zeckendorf representations. *Fibonacci Quart.*, 48(2) :168–174, 2010.

[15] Norbert Herrndorf. A functional central limit theorem for strongly mixing sequences of random variables. *Z. Wahrsch. Verw. Gebiete*, 69(4) :541–550, 1985.

[16] Xinxin Jiang and Marjorie Hahn. A self-normalized central limit theorem for ρ-mixing stationary sequences. *Statist. Probab. Lett.*, 78(12) :1541–1547, 2008.

[17] Sébastien Labbé and Jana Lepšová. A fibonacci's complement numeration system, 2022.

[18] C. G. Lekkerkerker. Representation of natural numbers as a sum of Fibonacci numbers. *Simon Stevin*, 29 :190–195, 1952.

[19] Johannes F. Morgenbesser and Lukas Spiegelhofer. A reverse order property of correlation measures of the sum-of-digits function. *Integers*, 12 :Paper No. A47, 5, 2012.

[20] Alexander Ostrowski. Bemerkungen zur Theorie der Diophantischen Approximationen. *Abh. Math. Sem. Univ. Hamburg*, 1(1) :77–98, 1922.

[21] Magda Peligrad. Invariance principles for mixing sequences of random variables. *Ann. Probab.*, 10(4) :968–981, 1982.

[22] Dimitris N. Politis, Joseph P. Romano, and Michael Wolf. *Subsampling*. Springer Series in Statistics. Springer-Verlag, New York, 1999.

[23] Emmanuel Rio. About the Lindeberg method for strongly mixing sequences. *ESAIM Probab. Statist.*, 1 :35–61, 1995/97.

[24] Emmanuel Rio. Inequalities and limit theorems for weakly dependent sequences. Lecture, September 2013.

[25] M. Rosenblatt. A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. U.S.A.*, 42 :43–47, 1956.

[26] Lukas Spiegelhofer. Correlations for numeration systems. Master's thesis, TU Wien and Aix-Marseille Université, 2014.

[27] Lukas Spiegelhofer. A lower bound for Cusick's conjecture on the digits of $n + t$. *Math. Proc. Cambridge Philos. Soc.*, 172(1) :139–161, 2022.

[28] Lukas Spiegelhofer and Michael Wallner. The digits of n+t. arXiv : 2005.07167, 2021.

[29] Jonas Kazys Sunklodas. Some estimates of the normal approximation for ϕ-mixing random variables. *Lith. Math. J.*, 51(2) :260–273, 2011.

[30] E. Zeckendorf. Représentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas. *Bull. Soc. Roy. Sci. Liège*, 41 :179–182, 1972.

Sur des propriétés stochastiques de la variation de la somme des chiffres en additionnant un entier fixé

Résumé. Le sujet de cette thèse vise à mieux comprendre les propriétés de la mesure de probabilité associée aux densités asymptotiques d'ensembles définis par la fonction somme des chiffres en base entière et non entière. Plus précisément, si on se donne un entier r, on s'intéresse aux ensembles d'entiers n tels que la différence entre la somme des chiffres de n et de $n+r$ est un entier relatif d fixé. La densité asymptotique $\mu^{(r)}(d)$ de ces ensembles peut être interprétée comme une mesure de probabilité sur les entiers. Évidemment, elle dépend du système de numération considéré.
Dans cette thèse, nous établissons des propriétés de la mesure de probabilité $\mu^{(r)}$, en base entière et en représentation de Zeckendorf. Pour cela, nous faisons le lien entre ces propriétés et certains systèmes dynamiques sous-jacents. En base entière, nos montrons notamment un théorème central limite. Nous généralisons certaines propriétés de $\mu^{(r)}$ en représentation de Zeckendorf et proposons un algorithme de calcul de $\mu^{(r)}$.

Mots clés : Théorème central limite, odomètre, fonction somme des chiffres, représentation de Zeckendorf, processus mélangeant.

On stochastic properties of the variation of the sum-of-digits function when adding a given integer

Abstract. The subject of this thesis aims to better understand the properties of the probability measure associated with asymptotic densities of sets defined by the digit sum in an integer base and not integer. More precisely, if we give ourselves an integer r, we are interested in sets of integers n such that the difference between the sum of the digits of n and $n+r$ is a fixed integer d. The asymptotic density $\mu^{(r)}(d)$ of these sets can be interpreted as a probability measure on the integers. Obviously, it depends on the numeration system we consider.
In this thesis, we describe properties of the probability measure $\mu^{(r)}$, in integer base and Zeckendorf representation. To do so, we make the link between these properties and some underlying dynamical systems. In integer base, we establish in particular, a central limit theorem. We generalise some properties of $\mu^{(r)}$ in Zeckendorf representation and give an algorithm to compute $\mu^{(r)}$.

Keywords: Central limit theorem, odometer, sum-of-digits function, Zeckendorf representation, mixing process.